TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

MASTERARBEIT

Function of the C-terminal regulatory domain of the cellulase and xylanase regulator XYR1 in *Trichoderma reesei*

ausgeführt am

Institut für Verfahrenstechnik, Umwelttechnik und technische Biowissenschaften der Technischen Universität Wien

unter Anleitung von

Privatdoz. Mag. Dr. rer. nat. Bernhard Seiboth und
Dipl.- Biol. Dr. Alexander Lichius (Ph.D)

durch

Franziska Buchholz
Preßgasse 13/5
1040 Wien

Wien, April 2014

DANKSAGUNG

An dieser Stelle möchte ich mich bei allen Menschen bedanken, die mich auf meinem bisherigen Weg tatkräftig unterstützt haben. Ich bedanke mich bei meinen Eltern, die mir in allen Bereichen und zu jeder Zeit zu Seite standen.

Ein großer Dank gilt allen Freunden im Besonderen meinem Freund Andreas und Judith, die mich während des Studiums und besonders neben der Diplomarbeit immer wieder auf andere Gedanken gebracht haben und mit Rat und Tat zur Seite standen.

Vielen Dank auch an alle meine KollegInnen für die nette Atmosphäre und die lustigen Momente im Labor. Ein besonders großer Dank gilt dabei vor allem Sara, Alexa und Tanja.

Ein herzliches Dankeschön an Alex für die kompetente Betreuung im Labor und beim Schreiben meiner Masterarbeit. Zum Schluss bedanke ich mich bei Bernhard für die Hilfe in allen Bereichen der Master- und Laborarbeit.

ZUSAMMENFASSUNG

Lignocellulose ist die größte erneuerbare Kohlenstoffquelle die heutzutage zur Verfügung steht und stellt somit eine wichtige Alternative zu Erdöl dar. Darum ist Lignocellulose wichtig für die Produktion von Bioethanol und dient ebenso als Rohstoff für Chemikalien. Die Umwandlung von Lignocellulose zu einzelnen Glukose Einheiten beinhaltet vor allem eine enzymatische Hydrolyse der Hauptkomponenten Cellulose und Hemicellulose. Ein weitverbreiteter Produzent von lignocellulolytischen Enzyme stellt der Pilz *Trichoderma reesei* dar. Die Herstellung der von diesem Pilz produzierte Enzymmischung ist jedoch sehr kostenintensiv und erfordert weitere Forschungen, um eine höhere Ausbeute in der Enzymproduktion zu erzielen.

Ein Weg auf der Suche nach neue Strategien für die Stammentwicklung stellt die vergleichende genomische Analyse von Stämmen mit einer starken Cellulase Produktion oder von Stämmen mit einer sehr geringen Cellulase Produktion zu den jeweiligen Elternstämmen dar. Eine genomische Analyse des Stammes QM9136 konnte eine Mutation identifizieren, die zu einem verkürzten XYR1 führt. In diesem verkürzte XYR1 fehlt die vermeintlich C-terminale regulatorische Domäne. Es zeigt sich darüber hinaus, dass beim Wiedereinsetzen eines Wild Typ *xyr1* der Phänotyp korrigiert werden konnte, sodass nun eine normale Cellulase Produktion möglich war. Dies lässt darauf schließen, dass die Mutation im *xyr1* zu einem Funktionsverlust führt.

Um diese Mutation auf molekularer Ebene untersuchen zu können, wurde von einer verkürzten Version des XYR1, das $XYR1^{1-780}$ N-terminal mit einem GFP fusioniert. So war es möglich den Transkriptionsfaktor XYR1 beim Shutteln zwischen dem Kern und dem Zytoplasma unter Cellulase induzierenden Bedingungen beobachtet zu können. Die Analyse zeigte eine basale nukleare Lokalisierung des GFP-$XYR1^{1-780}$. Jedoch konnte keine Erhöhung des GFP-$XYR1^{1-780}$ unter Cellulase induzierenden Bedingungen beobachtet werden im Vergleich zum nicht verkürzten XYR1, welches ebenfalls mit einem GFP N-terminal verbunden wurde (GFP-XYR1). Unter induzierenden Bedingungen zeigte sich, dass die *de-novo* Biosynthese von XYR1 sehr wichtig ist, um ausreichende Mengen an XYR1 zu produzieren. XYR1 wiederum wird in den Kern transportiert, um dort die Cellulase und Xylanase Transkription zu starten. Die verkürzte Version des XYR1, welches keine C-terminale regulatorische Domäne beinhaltet, verhindert die Hochregulierung von *xyr1* und kann durch die Autoregulation erklärt werden. Somit ist das $XYR1^{1-780}$ nicht in der Lage die Cellulase und Xylanase Gen Expression anzuregen.

ABSTRACT

Lignocellulosic biomass is the largest renewable carbon sources and is used today as important alternative to petroleum for the production of biofuels or platform chemicals. The bioconversion of lignocellulose to monomeric sugars involves also enzymatic hydrolysis of its main components cellulose and hemicellulose. The most common producer of lignocellulolytic enzymes is the fungus *Trichoderma reesei*. However, its enzyme cocktails are still not cost effective and further research is therefore necessary to improve the yields in enzyme production.

One way of finding new targets or strategies for strain improvement is the comparative genomic analysis of cellulase hyperproducing or non-producing strains to their parental strains. Genome analysis of QM9136 identified one mutation which resulted in a truncated XYR1 lacking the putative C-terminal regulatory domain. Reintroduction of the wild-type *xyr1* reversed the cellulase negative phenotype of QM9136 while introduction of the mutated version of *xyr1* in a *xyr1* deletion strain did not result in any phenotype, suggesting that the mutation in *xyr1* leads to a loss of function.

To investigate the molecular basis of the cellulase negative phenotype caused by the truncated version of XYR1, this XYR1$^{1-780}$ was N-terminal fused to GFP to follow its shuttling dynamics between cytoplasm and nucleus under cellulase inducing conditions. This analysis showed that a basal nuclear localization of GFP-XYR1$^{1-780}$ is found which is, however, not increased during under cellulase inducing conditions as found for a full length version of XYR1 fused to GFP (GFP-XYR1). Upon induction, *de-novo* protein synthesis of XYR1 is therefore essential for the generation of sufficient amounts of XYR1 which enter the nucleus to actively induce cellulase and xylanase transcription. The XYR1 truncation which lacks the C-terminal domain prevents XYR1 upregulation which can be explained by a XYR1 auto-regulation and is therefore also unable to stimulate cellulase and xylanase induction.

Table of Content

LIST OF ABBREVIATIONS

A. niger	*Aspergillus niger*
C_2H_2	Cys_2His_2
CBH	cellobiohydrolase
CBH1	cellobiohydrolase 1
CCR	carbon catabolite repression
CRE1	carbon catabolite repressor 1
EG	endoglucanase
GEF	guanine nucleotide exchange factor
NES	nuclear export signal
NLS	nuclear localization signal
T. reesei	*Trichoderma reesei*
TF	transcription factor
XYN2	xylanase 2
XYR1	xylanase regulator 1

1 Introduction

1.1 *Trichoderma reesei*

The filamentous fungus *Trichoderma reesei* was firstly discovered and isolated from the US Army on the Solomon Islands during the World War II. The reason for its discovery was its cellulolytic potential which was responsible for the degradation of the cotton-based army material of the US troops stationed there. Its name is based on the first principle investigator, Elwyn T. Reese (Reese 1976). *T. reesei* belongs to mesophilic ascomycetes, which are heterotrophic organisms. Many of them are capable of secreting large amounts of extracellular enzymes into their surroundings. These enzymes degrade complex polysaccharides that are for example, found in plant biomass, to their monomers. The resulting sugars are absorbed and provide the fungus with a carbon and energy source. *T. reesei* is widely used for the industrial production of cellulolytic and hemicellulolytic enzymes, which are for example, applied in the pulp and paper, or food and textile industries. Academic and industrial research programs classically employed random mutagenesis to improve enzyme yields in comparison to the original *T. reesei* isolate QM6a (Eveleigh et al. 1979; Durand et al. 1988).

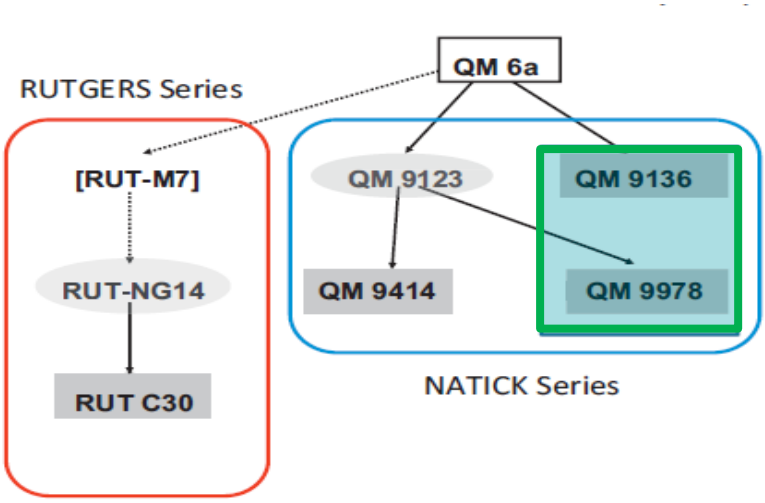


Fig. 1: Strain pedigree of *T. reesei* high and non-cellulase producers. Strains of the RUTGER series are shown on the left (framed in red); strains of the NATICK series are displayed on the right (framed in blue). QM 9123 and QM 9414 are early low cellulase and hemicellulase producers, whereas QM 9136 and QM 9987 are cellulase negative strains (framed in green). Diagram modified from (Kubicek 2013).

Fig. 1 shows that all mutated strains were derived QM6a the original wild type isolate of *T. reesei*. Strain QM6a has seven chromosomes and a genome size of 33Mb. Its genome was sequenced and published in 2008 (Martinez et al. 2008). Three isolates of the RUTGERS series were derived from QM6a; RUT C30 was produced from NG14 by UV- mutagenesis. RUT C30 has become the industrially-used reference strain among *T. reesei* high cellulase producer. The most prominent mutation found in this strain is the truncation of the carbon catabolite repressor CRE1 which leaves this strain carbon catabolite derepressed (Kubicek 2013). In contrast, the mutagenesis program of the NATICK mutagenesis program developed strains which became widely used in academic research (Reese 1976). Notably, strains QM9136 and QM 9978 turned out to be negative cellulase producers, whereas QM9414, derived from QM9123, was identified as an early positive low cellulase producer (Kubicek 2013).

1.2 Industrial importance of ligno-cellulosic biomass

Ligno-cellulosic biomass is an important element for bioethanol production. This biomass contains varying concentration of cellulose, hemicellulose and lignin and is an important part of plant cell walls. The annual production rate is in the range of about $7.2+6x10^{10}$ tons (Reese et al. 1950). Therefore ligno-cellulosic biomass represents one of the largest reservoirs for renewable carbon sources and thus an important alternative to petroleum based fuels. A great number of geopolitical aspects, such as environmental problems as a result of increased CO_2 emissions, climate changes and fossil fuels limitations are forcing investigations to implement and exploit renewable energy sources for the near future (Seidl et al. 2010). One of the biggest problems are the pollution emissions which are anthropogenic and loads the atmosphere with greenhouse gases, including carbon dioxide (CO_2), nitrous oxide (N_2O) and methane. The reason for the release of CO_2 is the burning of fossil carbon sources including oil, coal and natural gas. Because of the pollution of the environment, the shrinking deposits of fossil fuels, and constantly rising prices for energy, it is important to think about alternatives for a safe and sustainable energy. Efforts are made to replace fossil energy carriers with environmental friendly solar, wind or water energy (Shapouri et al. 2006). To reduce total carbon dioxide emissions, the use of biofuels and carbon dioxide recycling will be necessary. Therefore, feasible technologies for biofuel production are urgently required (Seidl et al. 2010).

Currently, the production of monomeric sugars from ligno-cellulosic biomass involves thermochemical process and enzymatic hydrolysis, before fermentation of sugar to ethanol can follow. Enzymatic hydrolysis is preferred for disintegration of the raw material because

thermochemical hydrolysis is very energy intensive (500-900°C) and results in the formation of inhibitory by-products (Shapouri et al. 2006). Therefore, enzymatic hydrolysis of ligno-cellulosic material, which operates at ambient temperatures (40°C to 50°C) represents a compatible method to gain carbon-neutral energy (Kubicek 2013). Dependent on the origin of the carbon source we distinguish between first and second generation biofuels: sugar or starch-based biomass constitutes the first generation of biofuels. The carbon sources are derived from food and are in general easily converted to biofuels. Sugarcane is frequently used in warmer regions, such as the world leader of first generation bioethanol production Brazil while sugar beets, wheat or corn are used in temperate regions, including Europe and the USA. However, both compete with the food and animal feed production and their agricultural production requires high energy input and the use of fertilizers. This greatly limits the reduction of total carbon dioxide emissions. Alternatively, second generation bioethanol is derived from non-food (parts of) plants, energy crops or ligno-cellulosic waste material. Energy crops can grow on marginal land, and reduce thereby the dependence of fertilizers. Second generation bioethanol production from ligno-cellulosic biomass therefore offers a future-proof strategy to significantly reduce carbon dioxide emissions for the production of renewable energy (Merino et al. 2007).

Nevertheless, improvements in different areas, including optimization of plant biomass pretreatment, genetic modification of the crops and optimization of enzyme mixtures and fermentation processes are necessary. Improvement of enzyme mixtures is the most critical part because high-performance enzyme cocktails which achieve high yield degradation of the plant material are currently too expensive to make industrial scale production economically feasible (Seidl et al. 2010). Hence, the costs for biofuel production from ligno-cellulosic material are not yet cost competitive to sugar- or starch-based biomass. *T. reesei* is the main producer of industrially-relevant plant cell wall degrading enzymes, improved strains secrete over 100 g/l cellulases (Cherry et al. 2003), which explains its outstanding value for this strategy.

1.3 Biopolymer structure and its degradation

Plant cell walls consist of different structural layers: the middle lamella, the primary cell wall and the secondary cell wall. Each layer consists of different biopolymers, comprising cellulose, hemicellulose and lignin. The secondary wall is located between the primary cell wall and the plasma membrane and contains mainly lignin and cellulose. It is built after completion of the

primary cell wall and after the cell has finished its expansion (Seidl et al. 2010; Kubicek et al. 2010; Wiley).

As it can be seen in Fig. 2, the β-(1,4)-linked glucose polysaccharide cellulose provides the structure and rigidity to the plant cell. This is also the reason for their resistance to chemical and enzymatic degradation. The multiple hydroxyl groups of the glucose monomers are linked to oxygen molecules on neighboring chains via hydrogen bounds. These bounds are responsible for the formation of microfibrils. The structural properties of cellulose depend on its chain length and degree of polymerization. The presence of lignin and hemicelluloses prevents an efficient enzymatic degradation by cellulases and the additional crystalline structure of the cellulose limits its accessibility (Seiboth et al. 2011).

Hemicelluloses are easily hydrolyzed by acids into their monomeric components, which comprise xylose, mannose, glucose, galactose, arabinose and certain amounts of rhamnose (Khandeparker et al. 2008). The outermost layer of the plant cell wall, the middle lamella, is mainly composed of pectin and has the role of connecting cells. Furthermore, pectin forms a gelatinous matrix within the primary cell wall, composed of homogalacturonan, rhamnogalactorunan I and rhamnogalactorunan II, in which the cellulose-hemicellulose network is embedded (Kubicek et al. 2010; Wiley). The final important biopolymer is lignin which provides structural integrity to plants (Khandeparker et al. 2008).

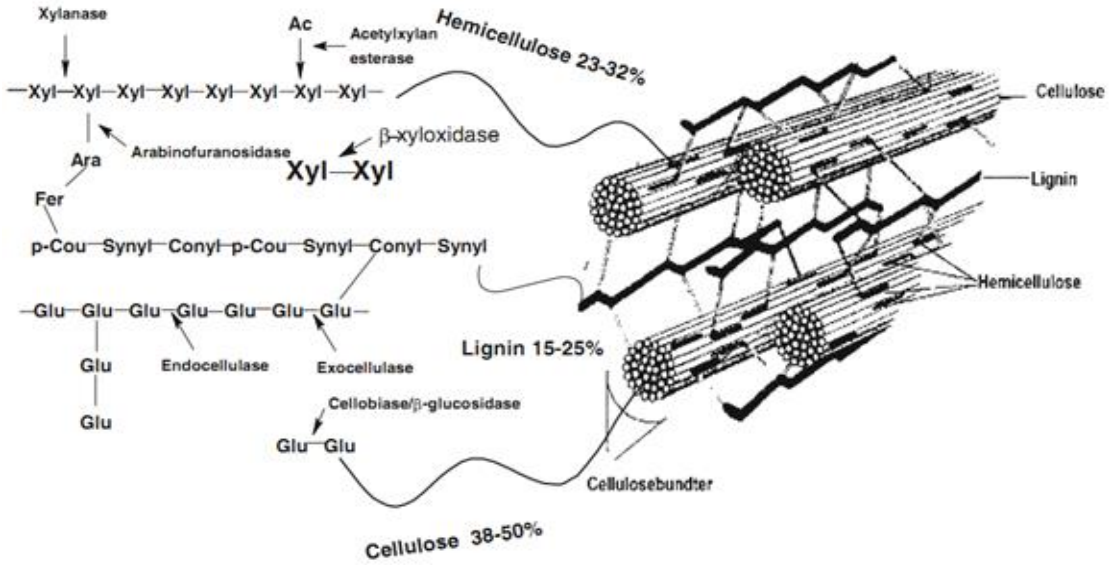


Fig. 2: The complex composition and structure of lignocelluloses require multiple enzymes for breaking down the polymers into sugar monomers. Cellulose is degraded by endocellulase, exocellulase, and ß–glucosidase. Figure taken from (Khandeparker et al. 2008).

Fig. 2 illustrates that different cellulases and hemicellulases are required for the complete degradation of lignocellulose fibers. In *T. reesei*, the cellulolytic system can be divided into three major enzyme classes. Some of the enzymes belonging originally to glycoside hydrolase family 61 were reclassified to the family of the lytic polysaccharide monooxygenases (Horn et al. 2012b). These oxidative enzymes produce also monomeric and dimeric oxidized sugars as part of the degradation process (Horn et al. 2012b).

Endoglucanases (EGs): degrade at random positions within the cellulose chain (Payne 2011; Horn et al. 2012).

Exoglucanases: they catalyze the cleavage of mainly cellobiose subunits of the reducing or non-reducing ends of the polymer and liberate the D-glucose dimer cellobiose from the ends of the cellulose chain. They are therefore also called cellobiohydrolases (CBHs) (Seiboth et al. 2011; Wiley).

β-glucosidases: release D-glucose from the soluble oligomeric breakdown products (Seiboth et al. 2011).

The majority of *T. reesei* enzymes comprise two cellobiohydrolases, CEL7A and CEL6A, which cleave the cellulose polymer in an exo-acting manner releasing mainly cellobiose from its reducing and non-reducing end, respectively. Both enzymes are composed of two structurally and functionally domains, including a catalytic domain and a carbohydrate binding module (CBMs) connected by a linker region. EGs have either a single catalytic domain or similar to the CBHs an additional carbohydrate binding module (CBM) (Seiboth et al. 2011).

1.4 Cellulase and xylanase gene regulation

As mentioned before, *T. reesei* is one of the many saprobic fungi capable of efficiently degrading cell wall polysaccharides such as cellulose and hemicelluloses. The genome of *T. reesei* encodes 10 cellulases and 16 hemicellulases (Martinez et al. 2008). The expression of cellulase and hemicellulase genes is tightly regulated as the biosynthesis of proteins and its secretion requires a lot of energy. Therefore, cell wall degrading enzymes are only synthesized in presences of an inducer, such as cellulose. Lactose is another very strong, however atypical inducer, and due to its low price and solubility of great importance for industrial production (Kubicek et al. 2009). The degradation product of cellulose, glucose is a strong repressor which

activates carbon catabolite repression (CCR). This mechanism allows the preferred assimilation of carbon sources of high nutritional values and is usually achieved through inhibition of gene expression of enzymes involved in the catabolism of other carbon sources or the uptake of other carbon sources (Portnoy et al. 2011b). Cellulase encoding genes are regulated by different specific transcription factors. Five of them have been identified: positive regulators are the cellulase and xylanase regulator 1 (XYR1), ACE2 and HAP2/3/5 complex; negative regulators comprise ACE1 and the carbon catabolite repressor 1 (CRE1) (Kubicek et al. 2009)

XYR1, the xylanase regulator 1, is considered to be the main activator of cellulase and hemicellulase gene expression and belongs to the Zn_2C_6 proteins (Stricker et al. 2006). Its genetic deletion resulted in elimination of cellulase induction by all known inducers and also induction of different hemicellulose genes which are involved in xylanase and arabinose degradation (Stricker et al. 2008). XYR1 and ACE2 are able to bind to the DNA binding domain motif [$GGAC(T)_4$]. The HAP 2/3/5 complex seems to be necessary for generation of an open chromatin structure and for full transcriptional activation (Zeilinger S. 2003).

On the other hand, ACE1 is able to repress XYR1 gene expression during growth on D-xylose (Aro et al. 2003). Elimination of the second cellulase activator, ACE2, seems to lower the transcript levels of the major cellulases and reduced cellulase activity from 30% to 70% when the fungus grows on cellulose, but showed no effects on cellulase induction by sophorose (Aro et al. 2001). In addition, XYR1 expression was elevated in a Δ*ace1* background, indicating that not only CRE1, but also ACE1 acts as a repressor of XYR1. It has been proposed, that the underlying ACE1 repression mechanism can be classified as “double-double” lock system (Portnoy et al. 2011).

As mentioned earlier, CRE1 belongs to the C_2H_2 type transcription factors and suppresses cellulase and hemicellulase gene expression by binding to the consensus sequence 5’-SYGGRG-3’ (Ilmén et al. 1996). It is considered to be the main transcription factor mediating CCR. CRE1 is related to the Mig1 protein that mediates glucose repression in *Saccharomyces cerevisiae* (Portnoy et al. 2011b).

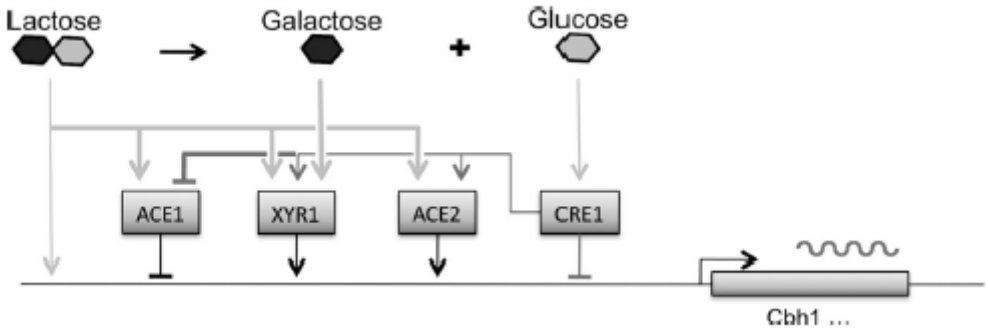


Fig. 3: Regulatory network of cellulase gene expression on lactose. Upon lactose induction, XYR1 and ACE2 are activated and *cbh*1 transcription starts. Glucose represses gene expression of XYR1 and ACE2 and activates CRE1 which is responsible for CCR. Taken from (Portnoy et al. 2011).

Fig. 3 schematically summarized the transcriptional regulation of the cellobiohydrolase 1 (*cbh*1) gene. It could be observed that the expression of the co-activator gene *ace*2 resembles that *xyr*1 in being inducible by lactose and also requiring CRE1 for a full induction. On the other hand side, glucose activates CRE1 which represses *chb*1 transcription as well as the expression of XYR1 and ACE1. It could be also shown that the expression of *ace*1 and its induction by lactose are repressed by CRE1 (Portnoy et al. 2011).

Beside the induction of cellulase, induction of xylanases is also important to consider. So far, five xylanases (XYN1-5) and have been identified in *T. reesei* including the recently described XYN5 (Herold et al. 2013). All xylanases are regulated via the transcriptional activator XYR1. Gene expression of *xyn1* is induced by D-xylose and gene expression by *xyn2* is in addition induced by xylobiose and the cellulase-inducing carbohydrates cellulose and sophorose (Zeilinger et al. 1997). *Xyn3* expression was originally only found in a mutant strain and is inducible by different cellulase inducers but not by D-xylose like the other xylanases. In *T. reesei*, *xyn1* and *xyn2* respond to CCR in different ways (Mach et al. 1996). L-arabinose is another inducer of xylanases and induces xylanase transcription independently of D-xylose (Herold et al. 2013).

1.5 Zinc cluster proteins

The trace element zinc is an important co-factor for various enzymes. Most of the zinc-containing proteins are transcription factors (TF) and, due to a specific structural DNA-binding motif, are also referred to as “zinc finger” proteins. Zinc-finger proteins form one of the largest

families of transcriptional regulators in eukaryotes. Cys-X_2-Cys-X_6-Cys-$X_{5–12}$-Cys-X_2-Cys-$X_{6–8}$-Cys is a fungal-specific zinc-binding motif (MacPherson et al. 2006). This zinc finger motif contains one α- helix and a pair of antiparallel β- strands in the 3D model which associates with e.g. CCG triplets in the DNA (Wolfe et al. 2000; Laity et al. 2001). One or more zinc atoms are bound by cysteine or histidine residues. So the domain is stabilized and contributes proper protein structures and functions. The majority of the zinc finger proteins play important roles in transcriptional and translational processes, including transcriptional control, RNA processing control, RNA transport control or translation control. Newly identified zinc finger proteins are also involved in many other physiological roles, such as protein-protein interactions, chromatin remodeling, protein chaperoning, lipid binding and zinc sensing (Laity et al. 2001).

Because of the different DNA-binding activities, three major classes of zinc-finger proteins can be distinguished in eukaryotes, based on their unique but highly conserved amino acid sequences (MacPherson et al. 2006): Class I zinc-finger TFs contains the Cys_2His_2 (C_2H_2) protein, the most common type of transcription factor (Wolfe et al. 2000). Class II represents the Cys_4 (C_4) zinc-finger TFs, which bind to a DNA sequence in the regulatory region of their target genes through zing-finger domains. Unlike the first class, these proteins usually contain one zinc finger unit binding to DNA as homodimers or heterodimers. Homodimers are supposed to invert repeats within the target nucleic acid sequence, whereas heterodimers bind to direct repeats (Laity et al. 2001). Class III zinc finger proteins contain a DNA-binding domain (DBD) which is important in this study. This protein class consist of six cysteine residues bound to two zinc atoms and is also referred to as zinc-binuclear-cluster, or $Zn(II)_2Cys_6$ (Zn_2Cys_6) TF. In this class one zinc finger unit binds to two zinc atoms. They seem to interact with DNA as monomers, homodimers, or heterodimers (MacPherson et al. 2006).

1.5.1 Structural and functional domains

In addition to the cysteine-rich DBD, zinc cluster proteins contain several other important functional domains apart from the cysteine-rich DBD. The region DBD itself is divided into three regions: the zinc finger itself, a linker and dimerization domain (Fig. 4).

Fig. 4: Functional protein domains of zinc cluster proteins. Zinc cluster proteins contain three structural domains: the DBD, the regulatory domain, and the acidic region. The DBD, furthermore, comprises three sub-regions: the zinc finger, the linker, and the dimerization domain. Taken from (MacPherson et al. 2006).

The metal-binding part of the DBD is has two substructures formed by three cysteine that are surrounded on both sides with basic amino acids and are separated by a loop (Schjerling et al. 1996). The linker region is located C-terminally to the zinc cluster motif and it can appear in different forms. Sequence alignments of various zinc cluster proteins revealed that these linkers are less conserved, and have thus been proposed to mediate sequence-specific DNA binding (MacPherson et al. 2006). The dimerization region is located C-terminally to the linker. This region is made up of heptad repeats and can be found in most, but not all, zinc proteins (Schjerling et al. 1996). These heptad repeats form a highly conserved coiled-coiled structure which is most likely responsible for dimerization and protein-protein interactions (MacPherson et al. 2006). The regulatory domain contains an important region and it is termed the middle homology region (MHR). It separates the DBD from the C-terminal acidic region and spans about 80 amino acids. It plays a role in regulating the transcriptional activity of the TF. Deletion of the MHR does not render the zinc cluster proteins inactive, suggesting that the C-terminally located, the acidic domain acts as the major activation domain (Schjerling et al. 1996; Turcotte et al. 2004). The acidic domain is not a highly conserved region and not well defined. Nevertheless, it has been shown that a deletion of about 10 C-terminal amino acids results in the inactivation of the transcription factor (MacPherson et al. 2006).

1.5.2 Nuclear import and localization

Zinc cluster proteins take part in transcriptional regulation but, as all proteins, are initially synthesized and often localized in the cytoplasm. Therefore, they must be imported into the nucleus. Nuclear import of transcription factors in eukaryotes occur by binding to a complex with an actively transport protein. The transport of these proteins is mediated through nuclear pores whereby soluble transport receptors bind to nuclear localization signals (NLSs) on their target molecules (Görlich et al. 1995; Görlich et al. 1999; Nakielny et al. 1999). Therefore,

proteins which were shuttled into the nucleus were bound by karyopherins the α/β importin heterodimer. The α subunit acts as a bridge between the NLS-containing cargo protein and the β subunit, which carries the cargo through the nuclear pore (MacPherson et al. 2006). Nuclear export is conferred by exportins that recognize nuclear export signals (NES). The motives are: $L(X)_{2-3}L(X)_{2-3}LXL$ (Bedard et al. 2007). The export is also an important regulatory mechanism by which the steady state level of a protein within the nucleus can be controlled (Beals et al. 1997; Gorner et al. 1998; Yan et al. 1998). The master regulator of nuclear transport is the Ras-like GTPase Ran (Melchior et al. 1993; Moore et al. 1993), which is thought to serve as a marker for the nuclear compartment (Izaurralde et al. 1997; Görlich 1998). Ran itself is present in the nucleus and the cytoplasm. The proteins that regulate its transition from inactive to active state by GDP for GTP exchange are compartmentalized. The Ran GTPase activating protein (GAP) is cytoplasmic whereas the guanine nucleotide exchange factor (GEF) is nuclear (Hopper et al. 1990; Clark et al. 1991; Becker et al. 1995). Inside the nucleus, Ran predominantly exists in the GTP-bound, active form, whereas in the cytoplasm the GDP-bound, inactive form dominates (Yan et al. 1998). The targets of the Ran are the β-karyopherins (β-importins). As described above, the karyopherins shuttle across the nuclear pore complex by binding to specific nucleoporins (Izaurralde et al. 1998). Because RanGTP is restricted to the nucleus, cargo should therefore be released preferentially within the nuclear interior, resulting in import against its own concentration gradient (Yan et al. 1998). This is crucial for achieving high level expression of target genes. Activation of zinc cluster proteins is also often mediated by phosphorylation. In addition, transcriptional activity is co-regulated by chromatin remodellers and histone modifiers (MacPherson et al. 2006).

1.6 Functional analysis of XYR1

The zinc binuclear transcriptional activator XYR1 in *T. reesei* activates the transcription of gene encoding cellulose- and hemicellulose- degrading enzymes (Stricker et al. 2008). The ortholog of *T. reesei* XYR1 in *Aspergillus niger* is the xylanolytic transcriptional activator (XlnR) (Hasper et al. 2004). Protein alignment of both transcription factors reveals a considerable degree of amino acid sequence homology (49.6%) (Lichius, A., unpublished data). In the following, the putative key roles of XYR1 in cellulase and xylanase gene expression will be explained using the current knowledge about functional protein domain structure of *A. niger* XlnR as a model (Fig. 5).

The (Zn_2Cys_6) DNA- binding domain which is found at the N- terminus is followed by an Arg-Arg-Arg-Leu-Trp amino acid motif, a marker for a fungal specific transcription-factor domain. An analysis of this region showed the presence of a putative coiled-coil domain (Lupas et al. 1991). C-terminally to this DNA-binding domain, a second coiled-coil region was identified. This second coiled-coil region might be involved in nuclear import dynamics (Hasper et al. 2004).

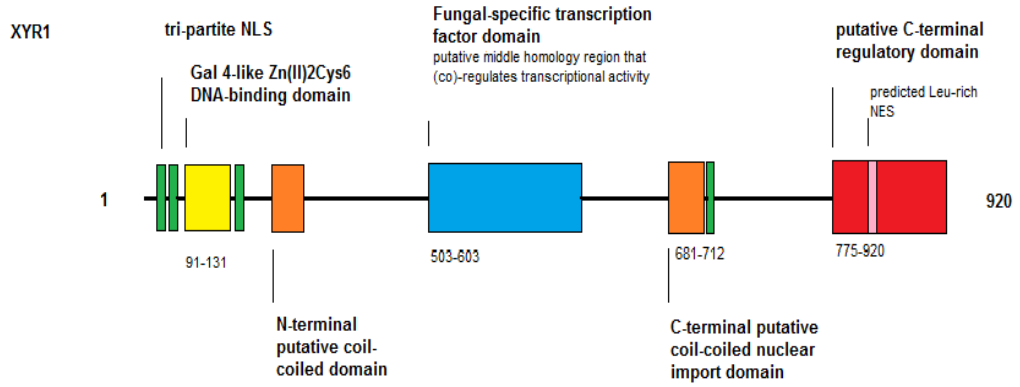


Fig. 5: XYR1 protein domain structure. The Zinc finger DNA-binding domain (yellow) is flanked by a tripartite NLS (Schneider-Poetsch et al.), followed by an N-terminal putative coil-coiled domain (orange). A fungal-specific transcription factor domain (blue) and C-terminal coil-coiled nuclear import domain with a mono-partite NLS (orange followed by green) is thought to be involved in nuclear import. The function of the putative C-terminal regulatory domain (red) is currently unknown. A predicted NES sequence is found in this region (pink).

Fig. 5 shows the respective orthologous domains of XYR1. The first functional domain at the N-terminus is the DNA-binding domain which is orthologous to the highly conserved transcription regulator Gal4 from *S. cerevisiae* (Schjerling et al. 1996). Gal4 is activated by phosphorylation, in the absence of galactose. Gal80 represses Gal4 by covering its activation (Ma et al. 1987). Under non-inducing conditions, an non-phosphorylated form and phosphorylated form are present (Sadowski et al. 1996). Hence, GaL4 binds to the promoter region of *gal* genes, but is blocked from activating transcription by the inhibitory protein Gal80. Induction by D-galactose needs the action of the signal transducer protein Gal3. The Gal3-Gal80 complex occurs in the cytoplasm and activates GAL4 in the nucleus (Blank et al. 1997; Peng et al. 2002).

Nuclear import of XYR1 appears to require a tripartite NLS, which can be identified located to either side of the Gal4 binding domain. There may also be a second putative coiled-coil domain involved in nuclear import, located towards the C-terminus. In *A. niger*, however, this putative nuclear import domain has been reported to be non-functional (Hasper et al. 2004). Notably, NLS, NES and DNA-binding domains of the Zinc finger cluster overlap which is a common

finding among zinc finger proteins (Fernandes- Martinez et al. 2003; Xiao et al. 2003; Ware et al. 2006). Previous studies have indicated the difficulty of differentiating import and export motif from DNA-binding domains due to functional overlap (LaCasse et al. 1995; Ware et al. 2006) It has been proposed that NLS and the DNA-binding domain have co-evolved in a way that active nuclear entry and DNA binding can occur independently (LaCasse et al. 1995). The overlap of these motifs may aid to release a protein from cognate importers (Bedard et al. 2007).

The next region is the fungal-specific transcription factor domain. This domain co-regulates transcriptional activity as described in more detail in section 1.5.1. (Schjerling et al. 1996; Turcotte et al. 2004). XYR1 has a C-terminal putative coiled-coil domain, which has been found in other transcriptional activators to be involved in dimerization (Hasper et al. 2004). The last domain region of XYR1 is the putative C-terminal regulatory domain which is conserved within the Zn_2Cys_6 transcription factor family. This domain contains a predicted nuclear export signal (NES) motif, characterized by hydrophobic and leucin-rich sequence (Fornerod et al. 1997; Fukuda et al. 1997). Mutants with a deletion of about 120 amino acids of the XlnR C-terminus downstream of the putative coiled-coil domain did not affect XlnR activity (Hasper et al. 2004)

In summary, the *in silico* comparison of the functional domain architecture of XlnR and XYR1 revealed many similarities between both. Important differences, however, have been found with respect to the regulation of cellulases and xylanases on the main inducers of the hemicellulolytic network. The cellulase and hemicellulase expression in *A. niger* is co-regulated via D-xylose, whereas in *T. reesei* at least three different inducer molecules (D-xylose, lactose and sophorose) are known. However, none of them alone seems capable of inducing all of the main cellulases and hemicellulases at the same time in *T. reesei* (Stricker et al. 2008). Another notable difference between *A. niger* and *T. reesei* is that D-xylose induces only xylanases but not cellulases in *T. reesei* (Ilmen et al. 1997). This seems to implicate that the substrate-unspecific activator XYR1 is fine-tuned by more specific transcriptional regulators (Seiboth et al. 2011).

1.7 Basic findings of XYR1 nucleo-cytoplasmic shuttling dynamics

Prior work by Lichius et al. (2014) established the basics for the study of the shuttling of XYR1 between nucleus and cytoplasm. To visualize GFP shuttling XYR1 was tagged with the green fluorescent protein (GFP)(Lichius et al. 2014). Only N-terminal GFP-labelling of XYR1, in contrast to attaching the GFP to the C-terminus, turned out to be essential for maintaining full functionality of the transcription factor. GFP-XYR1 showed a higher gene expression levels and improved efficiency in nuclear recruitment compared to the XYR1-GFP construct (Lichius et al. 2014). Testing various inducing and non-inducing conditions, it was established that the GFP-XYR1 fusion construct behaved similar to the XYR1 native protein. Nuclear localization of GFP-XYR1 specifically increased under cellulase-inducing conditions while nuclear export occurred quicker the stronger the non-inducing or repressing signal was. This work also revealed that cytoplasmic fluorescence intensity remained constant while the nuclear signal increased constantly during the induction period. This suggested that newly produced XYR1 became imported into the nuclei upon its biosynthesis and was not, as seen for CRE1, maintained mainly as cytoplasmic pool under non-inducing conditions. Export of XYR1 triggered by transfer to non-inducing conditions was, compared to CRE1 very slow. During this period the cytoplasm pool of XYR1 did not increase either, indicating rapid degradation of exported XYR1 (Lichius et al. 2014).

Fig. 6 schematically illustrates how TF shuttling dynamics and gene expression are functionally linked. That XYR1 nuclear import dynamics directly correlated with the rate of cellulase and xylanase gene expression was an important finding because nuclear presence of a transcription factor does not automatically imply that it is transcriptionally active.

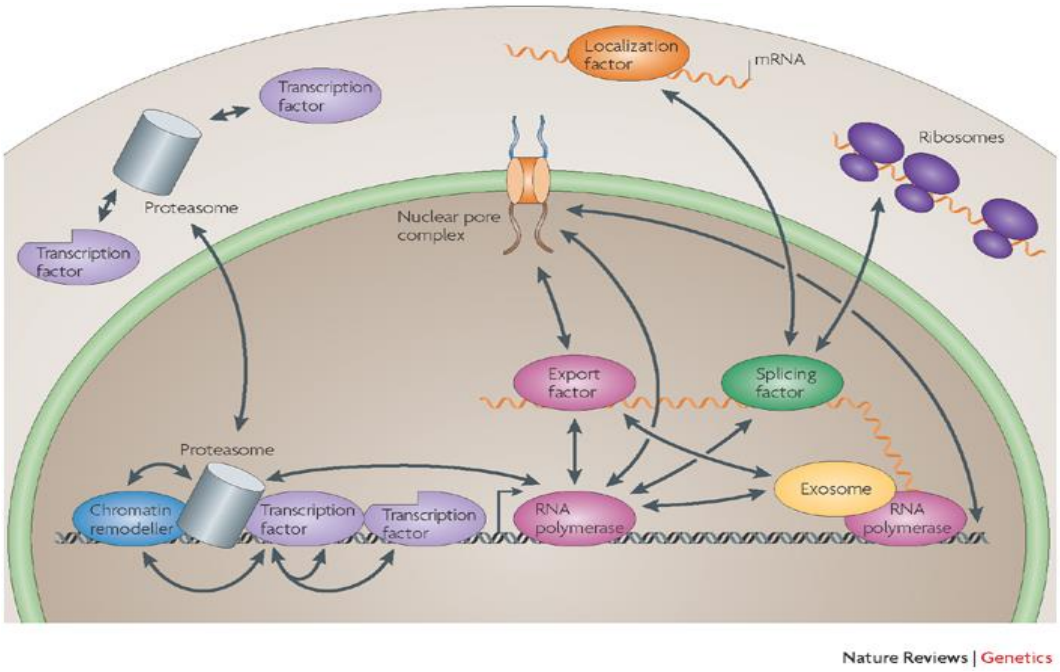


Fig. 6: Schematic representation of transcription factor shuttling with correlated gene expression. Upon induction TF are shuttled through the nuclear pore complex into the nucleus, where they bind to DNA, aided by other DNA-modifying proteins, such as chromatin remodellers. RNA polymerase becomes recruited and transcribes the activated gene region; here *cbh*1. The mRNA transcript is exported into the cytoplasm for ribosomal translation. Hence, quantification of isolated mRNA provides a reliable measure for specific gene activity. Taken from (www.nature.com)

Nuclear accumulation of GFP-XYR1 activated by the addition of sophorose (1.4mM), correlated with a strong *xyr1* and *cbh1* upregulation, but only very weak *xyn2* gene expression. The use of 1 mM xylose as inducing signal resulted in slower and less intense GFP-XYR1 nuclear import compared to sophorose correlating with weaker *cbh1* and *xyr1* upregulation, however, a significantly higher xylanase expression. Hence, kinetic differences in GFP-XYR1 shuttling dynamics truly reflected distinct metabolic cellular activity specific for the provided carbon source (Lichius et al. 2014).

Another key insight from these investigations was that XYR1-mediated gene regulation depends on its own *de-novo* biosynthesis. This was based on results from shuttling experiments under cellulase inducing conditions in combination with the protein biosynthesis blocker cycloheximide (Fig. 7). Upon parallel addition of sophorose and cycloheximide, only small amounts of pre-formed XYR1 shuttled into the nucleus which could not activate cellulase expression. Due to the protein synthesis inhibition by cycloheximide no further XYR1 protein was produced and an activation failed (Lichius et al. 2014).

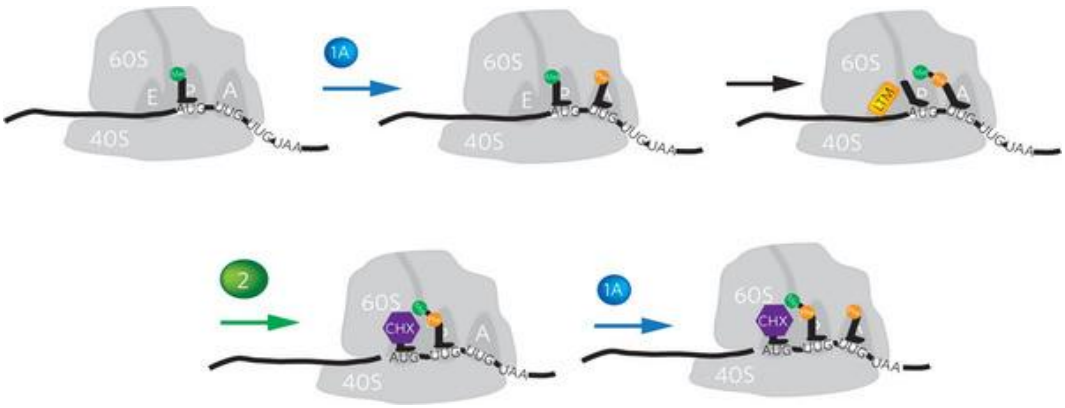


Fig. 7: Schematic representation of the steps of *de-novo* protein biosynthesis inhibition with cycloheximid. Cylcoheximid blocks the translation elongation step by preventing relative movement of tRNA and mRNA inside the ribosome (translation elongation freezing). Consequently, mRNA gene transcripts are still produced, but not translated into functional protein. Cycloheximide affects cellular protein biosynthesis except mitochondrial translation. Taken from (Schneider-Poetsch et al. 2010).

Although recent investigations widened our understanding on the molecular mechanism of XYR1-dependent cellulase and hemicellulase regulation, it is still not clear how XYR1 perceives the signal for nuclear import, and which nuclear carriers mediate its translocation. The present data showed that in the presence of glucose a basal nuclear localization of XYR1 exists, and is thus independent of inducing carbon sources. However, this basal XYR1 production is not sufficient for cellulase gene expression to occur, and probably maintained as a reservoir for the cell to be able to immediately respond to an inductive signal, or might be involved in other transcriptional activity. Upon cellulase-induction, *de-novo* protein synthesis of XYR1 is rapidly upregulated for the generation of sufficient amounts of XYR1 to enter the nucleus and be transcriptionally active for cellulase and xylanase production (Lichius et al. 2014).

1.8 QM9136 – a *T. reesei* cellulase negative mutant

Two of the negative cellulase producer strains from the NATICK series, QM9136 (Mandels et al. 1971) and QM9978 were sequenced and their mutations analyzed (unpublished data). The genome analysis of QM9136 showed one interesting mutation which resulted in a frame-shift in the coding region of *xyr1* resulting in several premature stop codons 5` of the putative C-terminal regulatory domain (Fig. 8). Therefore the resulting truncated XYR1 is the prime candidate to explain the loss-of-function phenotype of QM9136.

A

```
xyr1cDNA        ATGTTGTCCA ATCCTCTCCG TCGCTATTCT GCCTACCCCG ACATCTCCTC
Xyr1_9136       ATGTTGTCCA ATCCTCTCCG TCGCTATTCT GCCTACCCCG ACATCTCCTC

xyr1cDNA        GGCGTCATTT GACCCGAACT ACCATGGCTC ACAGTCGCAT CTCCACTCGA
Xyr1_9136       GGCGTCATTT GACCCGAACT ACCATGGCTC ACAGTCGCAT CTCCACTCGA

xyr1cDNA        TCAACGTCAA CACATTCGGC AACAGCCACC CCTATCCCAT GCAGCACCTC
Xyr1_9136       TCAACGTCAA CACATTCGGC AACAGCCACC CCTATCCCAT GCAGCACCTC

xyr1cDNA        GCACAGCATG CGGAGCTTTC GAGTTCACGC ATGATAAGGG CCAGTCCGGT
Xyr1_9136       GCACAGCATG CGGAGCTTTC GAGTTCACGC ATGATAAGGG CCAGTCCGGT

xyr1cDNA        GCAGCCAAAG CAGCGCCAGG GCTCTCTTAT TGCTGCCAGG AAGAATTCAA
Xyr1_9136       GCAGCCAAAG CAGCGCCAGG GCTCTCTTAT TGCTGCCAGG AAGAATTCAA

xyr1cDNA        CGGGTACTGC TGGGCCCATT CGGCGGAGGA TCAGTCGCGC TTGTGACCAG
Xyr1_9136       CGGGTACTGC TGGGCCCATT CGGCGGAGGA TCAGTCGCGC TTGTGACCAG

xyr1cDNA        TGCAACCAGC TTCGTACCAA GTGCGATGGC TTACACCCAT GTGCCCATTG
Xyr1_9136       TGCAACCAGC TTCGTACCAA GTGCGATGGC TTACACCCAT GTGCCCATTG

xyr1cDNA        TATAGAATTC GGCCTTGGAT GCGAATATGT CCGAGAGAGA AAGAAGCGTG
Xyr1_9136       TATAGAATTC GGCCTTGGAT GCGAATATGT CCGAGAGAGA AAGAAGCGTG

xyr1cDNA        GCAAAGCTTC GCGCAAGGAT ATTGCTGCCC AGCAAGCCGC GGCGGCTGCA
Xyr1_9136       GCAAAGCTTC GCGCAAGGAT ATTGCTGCCC AGCAAGCCGC GGCGGCTGCA

xyr1cDNA        GCACAACACT CCGGCCAGGT CCAGGATGGT CCAGAGGATC AACATCGCAA
Xyr1_9136       GCACAACACT CCGGCCAGGT CCAGGATGGT CCAGAGGATC AACATCGCAA

xyr1cDNA        ACTCTCACGC CAGCAAAGCG AATCTTCGCG TGGCAGCGCT GAGCTTGCCC
Xyr1_9136       ACTCTCACGC CAGCAAAGCG AATCTTCGCG TGGCAGCGCT GAGCTTGCCC

xyr1cDNA        AGCCTGCCCA CGACCCGCCT CATGGCCACA TTGAGGGCTC TGTCAGCTCC
Xyr1_9136       AGCCTGCCCA CGACCCGCCT CATGGCCACA TTGAGGGCTC TGTCAGCTCC

xyr1cDNA        TTCAGCGACA ATGGCCTTTC CCAGCATGCT GCCATGGGCG GCATGGATGG
Xyr1_9136       TTCAGCGACA ATGGCCTTTC CCAGCATGCT GCCATGGGCG GCATGGATGG

xyr1cDNA        CCTGGAAGAT CACCATGGCC ACGTCGGAGT TGATCCTGCC CTGGGCCGAA
Xyr1_9136       CCTGGAAGAT CACCATGGCC ACGTCGGAGT TGATCCTGCC CTGGGCCGAA

xyr1cDNA        CTCAGCTGGA AGCGTCATCA GCAATGGGCC TGGGCGCATA CGGTGAAGTC
Xyr1_9136       CTCAGCTGGA AGCGTCATCA GCAATGGGCC TGGGCGCATA CGGTGAAGTC

xyr1cDNA        CACCCCGGCT ATGAGAGCCC CGGCATGAAT GGCCATGTGA TGGTGCCCCC
Xyr1_9136       CACCCCGGCT ATGAGAGCCC CGGCATGAAT GGCCATGTGA TGGTGCCCCC

xyr1cDNA        GTCGTATGGC GCGCAGACCA CCATGGCCGG GTATTCCGGT ATCTCGTATG
Xyr1_9136       GTCGTATGGC GCGCAGACCA CCATGGCCGG GTATTCCGGT ATCTCGTATG

xyr1cDNA        CTGCGCAAGC CCCGAGTCCG GCTACGTATA GCAGCGACGG TAACTTTCGA
Xyr1_9136       CTGCGCAAGC CCCGAGTCCG GCTACGTATA GCAGCGACGG TAACTTTCGA

xyr1cDNA        CTCACCGGTC ACATCCATGA TTACCCGCTG GCAAATGGGA GCTCGCCCTC
Xyr1_9136       CTCACCGGTC ACATCCATGA TTACCCGCTG GCAAATGGGA GCTCGCCCTC

xyr1cDNA        ATGGGGACAA AGCGATTTGC GATATCCTGT GCTTGAGCCT CTGCTGCCTC
Xyr1_9136       ATGGGGACAA AGCGATTTGC GATATCCTGT GCTTGAGCCT CTGCTGCCTC

xyr1cDNA        ACCTGGGAAA CATCCTCCCC GTGTCTTTGG CGTGCGATCT GATTGACCTG
Xyr1_9136       ACCTGGGAAA CATCCTCCCC GTGTCTTTGG CGTGCGATCT GATTGACCTG

xyr1cDNA        TACTTCTCCT CGTCTTCATC AGCACAGATG CACCCAATGT CCCCATACGT
Xyr1_9136       TACTTCTCCT CGTCTTCATC AGCACAGATG CACCCAATGT CCCCATACGT
```

```
xyr1cDNA        TCTGGGCTTC GTCTTCCGGA AGCGCTCCTT CTTGCACCCC ACGAACCCAC
Xyr1_9136       TCTGGGCTTC GTCTTCCGGA AGCGCTCCTT CTTGCACCCC ACGAACCCAC

xyr1cDNA        GAAGGTGCCA GCCCGCGCTG CTTGCGAGCA TGCTGTGGGT GGCGGCACAG
Xyr1_9136       GAAGGTGCCA GCCCGCGCTG CTTGCGAGCA TGCTGTGGGT GGCGGCACAG

xyr1cDNA        ACTAGCGAAG CGTCCTTCTT GACGAGCCTG CCGTCGGCGA GGAGCAAGGT
Xyr1_9136       ACTAGCGAAG CGTCCTTCTT GACGAGCCTG CCGTCGGCGA GGAGCAAGGT

xyr1cDNA        CTGCCAGAAG CTGCTCGAGC TGACCGTTGG GCTTCTTCAG CCCCTGATCC
Xyr1_9136       CTGCCAGAAG CTGCTCGAGC TGACCGTTGG GCTTCTTCAG CCCCTGATCC

xyr1cDNA        ACACCGGCAC CAACAGCCCG TCTCCCAAGA CTAGCCCCGT CGTCGGTGCT
Xyr1_9136       ACACCGGCAC CAACAGCCCG TCTCCCAAGA CTAGCCCCGT CGTCGGTGCT

xyr1cDNA        GCTGCCCTGG GAGTTCTTGG GGTGGCCATG CCGGGCTCGC TGAACATGGA
Xyr1_9136       GCTGCCCTGG GAGTTCTTGG GGTGGCCATG CCGGGCTCGC TGAACATGGA

xyr1cDNA        TTCACTGGCC GGCGAAACGG GTGCTTTTGG GGCCATAGGG AGCCTTGACG
Xyr1_9136       TTCACTGGCC GGCGAAACGG GTGCTTTTGG GGCCATAGGG AGCCTTGACG

xyr1cDNA        ACGTCATCAC CTATGTGCAC CTCGCCACGG TCGTCTCGGC CAGCGAGTAC
Xyr1_9136       ACGTCATCAC CTATGTGCAC CTCGCCACGG TCGTCTCGGC CAGCGAGTAC

xyr1cDNA        AAGGGCGCCA GCCTGCGGTG GTGGGGTGCG GCATGGTCTC TCGCCAGAGA
Xyr1_9136       AAGGGCGCCA GCCTGCGGTG GTGGGGTGCG GCATGGTCTC TCGCCAGAGA

xyr1cDNA        GCTCAAGCTT GGCCGTGAGC TGCCGCCTGG CAATCCACCT GCCAACCAGG
Xyr1_9136       GCTCAAGCTT GGCCGTGAGC TGCCGCCTGG CAATCCACCT GCCAACCAGG

xyr1cDNA        AGGACGGCGA GGGCCTTAGC GAAGACGTGG ATGAGCACGA CTTGAACAGA
Xyr1_9136       AGGACGGCGA GGGCCTTAGC GAAGACGTGG ATGAGCACGA CTTGAACAGA

xyr1cDNA        AACAACACTC GCTTCGTGAC GGAAGAGGAG CGCGAAGAGC GACGGCGAGC
Xyr1_9136       AACAACACTC GCTTCGTGAC GGAAGAGGAG CGCGAAGAGC GACGGCGAGC

xyr1cDNA        ATGGTGGCTC GTTTACATCG TCGACAGGCA CCTGGCGCTC TGCTACAACC
Xyr1_9136       ATGGTGGCTC GTTTACATCG TCGACAGGCA CCTGGCGCTC TGCTACAACC

xyr1cDNA        GCCCCTTGTT TCTTCTGGAC AGCGAGTGCA GCGACTTGTA CCACCCGATG
Xyr1_9136       GCCCCTTGTT TCTTCTGGAC AGCGAGTGCA GCGACTTGTA CCACCCGATG

xyr1cDNA        GACGACATCA AGTGGCAGGC AGGCAAATTT CGCAGCCACG ATGCAGGGAA
Xyr1_9136       GACGACATCA AGTGGCAGGC AGGCAAATTT CGCAGCCACG ATGCAGGGAA

xyr1cDNA        CTCCAGCATC AACATCGATA GCTCCATGAC GGACGAGTTT GGCGATAGTC
Xyr1_9136       CTCCAGCATC AACATCGATA GCTCCATGAC GGACGAGTTT GGCGATAGTC

xyr1cDNA        CCCGGGCGGC TCGCGGCGCA CACTACGAGT GCCGCGGTCG TAGCATTTTT
Xyr1_9136       CCCGGGCGGC TCGCGGCGCA CACTACGAGT GCCGCGGTCG TAGCATTTTT

xyr1cDNA        GGCTACTTCT TGTCCTTGAT GACAATCCTG GGCGAGATTG TCGATGTCCA
Xyr1_9136       GGCTACTTCT TGTCCTTGAT GACAATCCTG GGCGAGATTG TCGATGTCCA

xyr1cDNA        CCATGCTAAA AGCCACCCCC GGTTCGGCGT TGGATTCCGC TCCGCGCGGG
Xyr1_9136       CCATGCTAAA AGCCACCCCC GGTTCGGCGT TGGATTCCGC TCCGCGCGGG

xyr1cDNA        ATTGGGACGA GCAGGTTGCT GAAATCACCC GACACCTGGA CATGTATGAG
Xyr1_9136       ATTGGGACGA GCAGGTTGCT GAAATCACCC GACACCTGGA CATGTATGAG

xyr1cDNA        GAGAGCCTCA AGAGGTTCGT GGCCAAGCAT CTGCCATTGT CCTCAAAGGA
Xyr1_9136       GAGAGCCTCA AGAGGTTCGT GGCCAAGCAT CTGCCATTGT CCTCAAAGGA

xyr1cDNA        CAAGGAGCAG CATGAGATGC ACGACAGTGG AGCGGTAACA GACATGCAAT
Xyr1_9136       CAAGGAGCAG CATGAGATGC ACGACAGTGG AGCGGTAACA GACATGCAAT

xyr1cDNA         CTCCACTCTC GGTGCGGACC AACGCGTCCA GCCGCATGAC GGAGAGCGAG
Xyr1_9136       CTCCACTCTC GGTGCGGACC AACGCGTCCA GCCGCATGAC GGAGAGCGAG
```

```
xyr1cDNA        ATCCAGGCCA GCATCGTGGT GGCTTACAGC ACCCATGTGA TGCATGTCCT
Xyr1_9136       ATCCAGGCCA GCATCGTGGT GGCTTACAGC ACCCATGTGA TGC.TGTCCT
                                                        missing 'A'

xyr1cDNA        CCACATCCTC CTTGCGGATA AGTGGGATCC CATCAACCTT CTAGACGACG
Xyr1_9136       CCACATCCTC CTTGCGGATA AGTGGGATCC CATCAACCTT CTAGACGACG
                                  premature stopps
xyr1cDNA        ACGACTTGTG GATCTCGTCG GAAGGATTCG TGACGGCGAC GAGCCACGCG
Xyr1_9136       ACGACTTGTG GATCTCGTCG GAAGGATTCG TGACGGCGAC GAGCCACGCG

xyr1cDNA        GTATCGGCTG CCGAAGCTAT TAGCCAGATT CTCGAGTTTG ACCCTGGCCT
Xyr1_9136       GTATCGGCTG CCGAAGCTAT TAGCCAGATT CTCGAGTTTG ACCCTGGCCT

xyr1cDNA        GGAGTTTATG CCATTCTTCT ACGGCGTCTA TCTCCTGCAG GGTTCCTTCC
Xyr1_9136       GGAGTTTATG CCATTCTTCT ACGGCGTCTA TCTCCTGCAG GGTTCCTTCC

xyr1cDNA        TCCTCCTGCT CATCGCCGAC AAGCTGCAGG CCGAAGCGTC TCCAAGCGTC
Xyr1_9136       TCCTCCTGCT CATCGCCGAC AAGCTGCAGG CCGAAGCGTC TCCAAGCGTC

xyr1cDNA        ATCAAGGCTT GCGAGACCAT TGTTAGGGCA CACGAAGCTT GCGTTGTGAC
Xyr1_9136       ATCAAGGCTT GCGAGACCAT TGTTAGGGCA CACGAAGCTT GCGTTGTGAC

xyr1cDNA        GCTGAGCACA GAGTATCAGC GCAACTTTAG CAAGGTTATG CGAAGCGCGC
Xyr1_9136       GCTGAGCACA GAGTATCAGC GCAACTTTAG CAAGGTTATG CGAAGCGCGC

xyr1cDNA        TGGCTCTGAT TCGGGGCCGT GTGCCGGAAG ATTTAGCTGA GCAGCAGCAG
Xyr1_9136       TGGCTCTGAT TCGGGGCCGT GTGCCGGAAG ATTTAGCTGA GCAGCAGCAG

xyr1cDNA        CGACGACGCG AGCTTCTTGC ACTATACCGA TGGACTGGTA ACGGAACCGG
Xyr1_9136       CGACGACGCG AGCTTCTTGC ACTATACCGA TGGACTGGTA ACGGAACCGG

xyr1cDNA        TCTGGCCCTC TAA
Xyr1_9136       TCTGGCCCTC TAA
```

B

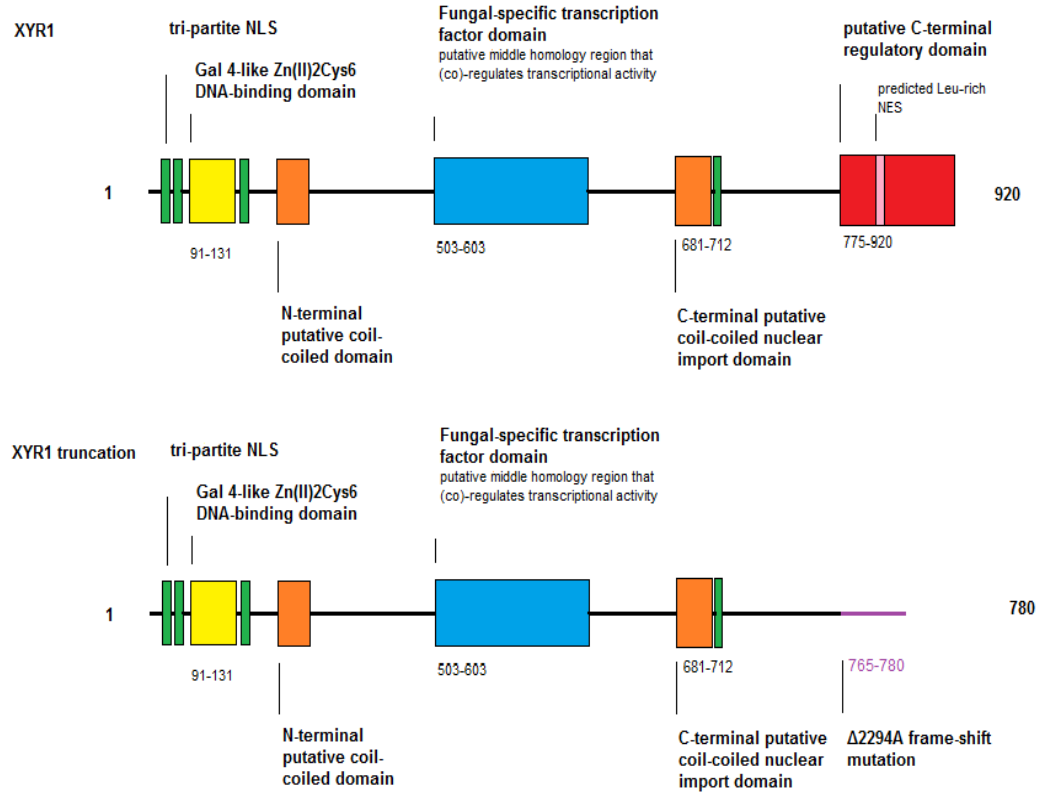


Fig. 8: Functional protein domain comparison between wild type and mutated XYR1. (A) Alignment of the native *xyr1* of QM6a and the mutant allele encoded in QM9136. **(B)** Protein domain comparison between wild-type XYR1 and truncated XYR1. At position 2294 (Δ2294A), the deletion of an adenosine leads to a frame-shift mutation that introduces a series of stop codons just in front of the putative regulatory domain. The first of these terminates translation prematurely after amino acid 780. This results in a truncated $XYR1^{1-780}$ protein whereby the last 16 are the result of the deletion and are not present in XYR1. $XYR1^{1-780}$ completely lacks the putative regulatory domain spanning from amino acid position 775 until 920.

A previous study analyzed the growth and enzyme production properties in submerged cultures of QM9136 compared to the wild type strain QM6a, and showed that the growth was rapid on starch and sugars (Mandels et al. 1971). Notably, QM9136 did not grow on cellulose, and isolated enzymes showed no cellulase activity *in vitro* (Mandels et al. 1971). Torigoi *et al.* (1996) confirmed that the mutations found in QM9136 and QM9978 specifically affected the expression of cellulose-induced genes. Based on a reporter gene study, they concluded that a basal level of cellulases is produced. Furthermore, mutants of *A. niger* lacking about 120 amino acids of the putative regulatory C-terminal domain of XlnR have been reported to be defective in nuclear import of the protein (Hasper et al. 2004). This suggests that the orthologous region

in *T. reesei* XYR1 probably is involved in nuclear import as well. Therefore, the functional analysis of the mutated XYR1$^{1-780}$ protein might reveal novel details on the mechanism underlying XYR1 shuttling and efficient cellulase production.

1.9 Aim of the study

The aim of this Masters project was to investigate the role of the putative regulatory C-terminal domain of XYR1 absent in strain QM9136 in nuclear import and transcriptional regulation during cellulase and xylanase induction in *T. reesei*.

Therefore the native *xyr1* locus in QM9414 will be replaced by the QM9136 truncated form of *xyr1* fused to the GFP gene resulting in GFP-XYR1$^{1-780}$ to allow live-cell imaging analysis of the nucleo-cytoplasmic shuttling dynamics. As controls, QM9136 and QM9414 strains transformed with the functional full-length GFP-XYR1 will be used.

In order to correlate nuclear presence of functional and non-functional XYR1 alleles with transcriptional activity, expression profiles of marker genes, including both transcription factors *xyr1* and *cre1*, and the cellulase and xylanase encoding genes *cbh1* and *xyn2*, respectively, were monitored using real time quantitative PCR (RT-qPCR).

This correlative live-cell imaging/gene expression analysis of a fluorescently labelled, truncated XYR1 construct shall then clarify the functional role of the XYR1 C-terminus with regard to the role of the C-terminal truncation in nuclear import and cellulase and xylanase gene expression.

2 Materials and Methods

2.1 Strains

2.1.1 Fungal strains

T. reesei strains QM9414 (ATCC 26921) is an early cellulase hyper-producing mutant and the cellulase negative strain QM9136 (ATCC 26920) were used throughout this work. Both derived from the wild type isolate QM6a (ATCC 269219). The gene deletion strains Δ*xyr1* and Δ*cre1* are also derived from QM9414. QM9414 and QM9414 Δ*tku70* respectively, served as reference strain in all experiments. *T. reesei* transformants were created in a QM9414 Δ*tku70* strain background, in the latter one the Δ*tku70* gene was replaced in strain QM9414 by the pyrithiamine marker (Kubodera et al. 2002). Strains expressing GFP-XYR1 fusion proteins comprised TRAL002 (GFP-XYR1), TRAL007 (GFP-XYR1$^{1-780}$) and TRAL008 (GFP-XYR1). The strains QM9136 + XYR1 and Δ*xyr1* + XYR1$^{1-780}$ were generated by Lukas Hartl (unpublished data). In QM9136 *xyr1* with a point mutation (Δ2294A) is expressed leading to the truncated XYR1$^{1-780}$. Introduction of the wild type *xyr1* from QM9414 into QM9136 complemented this transformant leading to normal growth on xylose. In the Δ*xyr1* strain the *xyr1* locus was exchanged with Pxyr1-*amdS*-Txyr1 removing the coding region of xyr1 (Stricker et al. 2006). Then Δ*xyr1* was complemented with a construct of Pxyr1- XYR1$^{1-780}$-Txyr1 (*xyr1*$^{\Delta 2294A}$) to generate the transformant Δ*xyr1* + XYR1*. All strains used are listed in Table 1.

Table 1: *T. reesei* strains which were used in this work.

Strain name	Protein	Parental strain	Genotype	Reference
QM9414 (ATCC 26921)	XYR1	QM6a	reference strain	(Reese 1976)
QM9414 Δ*tku70*	XYR1	QM9414	Δ*tku70*::*ptrA*	(Ivanova *et al*. unpublished data)
QM9136 (ATCC 26920)	XYR1$^{1-780}$	QM6a	frame-shift point mutation (Δ2294A) leading to truncation of XYR1 C-terminus	(Mandels et al. 1971)
Δxyr1	No XYR1	QM9414 *Δtku70*	Δ*xyr1*::*amdS*	(Stricker et al. 2006)

Δ*cre1*	XYR1, kein CRE1	QM9414	Δ*cre1*::*hph*	(Portnoy et al. 2011b)
QM9136 + XYR1 QM91.12, QM91.20, QM91.27	XYR1$^{1-780}$ + XYR1	QM9136	P*xyr1*::*xyr1*::T*xyr1*; ectopic integration	(Hartl, L. unpublished data)
Δ*xyr1* + XYR1 $^{\Delta 1-780}$ **(XYR26, XYR27)**	XYR1$^{1-780}$	Δ*xyr1*	P*xyr1*::*xyr1*$^{\Delta 2294A}$::T*xyr1* ; ectopic integration of *xyr1*$^{\Delta 2294A}$	(Hartl, L. unpublished data)
QM9414+GFP-XYR1 (TRAL002)	GFP-XYR1	QM9414 *Δtku70*	Δ*tku70*::*ptrA*; Δ*xyr1*::*hph*; P*xyr1*::*gfp-xyr1*::T*xyr1* (homologously integrated)	(Lichius et al. 2014)
QM9136+GFP-XYR1 (TRAL008)	GFP-XYR1	QM9136	Δ*xyr1*::*hph*; P*xyr1*::*gfp-xyr1*::T*xyr1* (homologously integrated)	(Lichius et al. 2014)
QM9414 + GFP-XYR1$^{1-780}$ (TRAL007)	**GFP-XYR1$^{\Delta 1-780}$**	QM9414 Δ*tku70*	Δ*tku70*::*ptrA*; Δ*xyr1::hph*; P*xyr1*::*gfp-xyr1*$^{\Delta 2294A}$::T*xyr1*, in locus	This study

2.1.2 Bacterial strains

Escherichia coli JM109 was used for plasmid constructions and amplification using standard molecular cloning techniques (Sambrook et al. 2006).

2.2 Media and solutions

2.2.1 Potato dextrose agar (PDA)

39g/l PDA (Difco) were dissolved in tap water, autoclaved and tempered to 50°C prior to pouring into Petri dishes.

2.2.2 Lysogeny broth medium

For *E. coli* cultivations, LB medium with the following composition was used (Fernandes-Martinez et al.) :

Peptone	10g/l
Yeast extract	5g/l
NaCl	10g/l

For plates 1.5% (w/v) agar- agar was added. After adjusting to the desired volume with distilled water, the medium was autoclaved and tempered to 50°C. Then 1µl/ml of a 1000x concentrated solution of ampicillin was added to the medium (LB amp Medium).

2.2.3 SOC medium

For transformation of *E. coli*, SOC medium which was used for a recovery step after heat-shock transformation (see 3.4.5) with the following composition was used:

Trypton	20g/l
Yeast extract	5g/l
NaCl 0.5	5g/l
0,25M KCl	10ml
2M $MgCl_2$	5ml
1M Glucose	20ml

The medium was filled up with distilled water up to a final volume of 1L. NaOH was added to a final pH value of 7.0.

2.2.4 Mandels Andreotti medium

For cultivation of fungal strains in liquid culture, minimal medium with the following composition was used:

KH_2PO_4	2g/l
$(NH_4)SO4$	1.4g/l
$CaCl_2*2H_2O$	0.3g/l
$MgSO_4*7H_2O$	0.3g/l
100x Trace Element Stock	10ml/l

100x Trace Element Stock:

$FeSO_4*7H_2O$	0.5g/l
$ZnSO_4*7H_2O$	0.2g/l
$MnSO_4*7H_2O$	0.2g/l
pH 5.5	
Peptone:	0.05%
Glucose:	1%

Minimal Liquid Medium was autoclaved and tempered to 50°C.

2.3 Culture conditions

2.3.1 Media and solutions for protoplast transformation

3.3.5.1. Bottom medium

PDA and 1M sorbitol were dissolved in distilled water, autoclaved and tempered to 50°C. For antibiotic selection, 100µg/µl of Hygromycin B (Calbiochem) was added

3.3.5.2. Overlay medium

1% (w/v) of Agar noble (Difco) and 1M sorbitol were dissolved in distilled water, autoclaved and tempered to 50°C. For antibiotic selection, 100µg/µl of Hygromycin B were added.

3.3.5.3. Solution A

1.2M sorbitol and 0.1M KH_2PO_4 were dissolved in distilled water. The solution was adjusted to pH= 5.6 and subsequently autoclaved.

3.3.5.4. Solution B

1M sorbitol, 50mM CaCl2 and 10mM Tris- HCl were dissolved in distilled water. The solution was adjusted to pH= 7.5 and then autoclaved.

3.3.5.5. PEG Solution

25% PEG 6000, 50mM $CaCl_2$ and 10mM Tris- HCl were dissolved in distilled water. The solution was adjusted to pH= 7.5 and thereupon autoclaved.

2.3.2 Medium for single spore isolation

In order to obtain purified colonies from single spores, transformants were transferred to PDA plates containing 0.1% Triton- X 100 and 100µg/µl of the antibiotic Hygromycin B.

2.3.3 Cultivation of fungal strains

All strains were maintained on PDA medium. Plates were incubated at 28°C until sporulation.

2.3.4 Cultivation of bacteria

The cultivation of *E. coli* for plasmid amplifications was carried out overnight at 37°C in LB medium containing ampicillin. Cultivations in liquid media were incubated in the water bath at 37°C with shaking (150rpm).

2.3.5 Growth tests of fungal strains on MA plates

For assessing the growth of *T. reesei* strains and the transformant generated in this study (see Table 2), 5µl spore solution of each of the respective strains were placed in the center of a PDA plate and incubated at 12h/ 12h light/ darkness cycles at 28°C. Measurement of radial growth was started 24 hours, carried out at 48 hours and 72 hour after incubation. Each of the respective strains was placed on different carbon source. Pictures were taken after 96 hours.

Table 2: Used carbon sources and concentrations.

Carbon source	Concentration [w/v]	Concentration [mM]
Glucose	1%	50
Lactose	1%	25
Xylose low	0.015%	1
Xylose high	1%	65
Carboxymethylcellulose	1%	-

2.3.6 Cultivation for real-time monitoring of transcription factor shuttling and gene expression analysis

Direct cultivation: For direct cultivation, 2 million spores per mL were inoculated in MA media containing 1% carbon source.

Carbon source replacement: For replacement experiments a fully sporulated MA plate was harvested in 15mL of NaCl Tween 80. In case strains sporulate poorly, longer incubation times were used (about 3-4 days) and spore clumps were resuspended by pipetting up and down. The obtained spore suspensions were filtered through a sterile Miracloth filter. Subsequently, 2 mL of the filtered spore suspension were inoculated in 100 mL MA and 1% glucose and incubated at 28°C and 250 RPM for 22 hours. 15 – 20ml mycelium containing culture medium were filtered and washed with sterile tap- water, finally and suspended in 20ml sterile replacement MA. After 10 min at 28°C and 250 rpm, the new carbon source was added. For life-cell imaging 25ml samples were drawn at desired times and analyzed on the confocal laser scanning microscope. For gene expression analysis, biomass from original and replaced cultures was harvested over sterile Miracloth, washed with ice cold water, dry-pressed and immediately frozen in liquid nitrogen. Until further processing they were stored at -80°C.

2.3.7 Harvesting of fungal spores

Carbon source free MA plates were centrally inoculated with 5µl of aqueous spores suspension from fridge stocks. The plates were incubated for 4-5 days at 28°C in the dark, until sufficient conidiation had occurred. Spores were harvested by rinsing the whole plate with NaCl/ Tween solution (0.9% NaCl, 0.05% Tween), and filtering through a sterile Miracloth-filter tip to remove hyphal fragments.

2.4 Molecular biology standard techniques

2.4.1 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is a molecular biological method to amplify a specific part of DNA sequence by repeated *in vitro* synthesis (Mullis 1990). The reaction mix contains:

1. DNA template (in this work: genomic DNA or plasmid DNA)
2. two oligonucleotides (primers) which bind at the 5' and 3' end of the fragment which is to be amplified
3. DNA Polymerase
4. Deoxynucleosidetriphosphates (dATP, dCTP, dGTP, dTTP)
5. Polymerase buffer

One complete run consists of three steps, denaturation, annealing and elongation:

Denaturation: Performed at 98°C to separate the DNA double helix to two single strands.

Annealing: In this step, the two primers bind to the sequence to be amplified from the template DNA. The annealing temperature is based on the melting temperature of primers and usually lies between 55°C and 65°C for 30 seconds.

Elongation: In this step the DNA polymerase starts to synthesize the complement strand in 5' to 3' direction starting at the 3' end of the annealed primer. The temperature and time of this step is based on the enzyme and the fragment length which is used. In this study PHIRE® and PHUSION® polymerase both had working temperatures of 72°C.

Usually the amplified fragment is checked for correct size following an agarose gel electrophoresis (0.8-2%). Dependent on the downstream process the amplified fragment can be column purified using commercial purification kits.

2.4.2 Agarose gel electrophoresis (DNA)

For the analysis of DNA fragments after PCR or endonuclease restriction digest, agarose gel electrophoresis is a commonly used technique. For this, 1g agarose was dissolved in 100ml of 1x TAE buffer microwave-heated until the solution was clear. After cooling down and adding 2.5µl of SYBR® Safe DNA Gel Stain (Invitrogen), the gel was poured into a gel tray, prepared with the appropriate comb to create pockets in the gel. Upon solidification of the gel matrix, the comb was removed and the gel tray was placed in an electrophoresis chamber filled with 1x TAE buffer. Thereafter, the gel pockets were loaded with 5µl of a molecular weight standard solution (50bp or 1kb GeneRulerTM DNA Ladder, Fermentas) and the samples, which were previously blended with loading dye. DNA samples were separated in the gel using a working voltage between 80 and 120V.

TAE buffer (50 x):

Tris (2M)	242 g/l
1M Acetic acid	57.1 ml
50mM EDTA	100 ml
ddH_2O	add 1l
Adjusted to pH 8	

2.4.3 Generation of gene replacement cassette

The replacement cassette for the expression of N-terminally tagged GFP-XYR1$^{1-780}$ -fusion construct from the native *xyr1* locus was cloned by modifying an existing pGFP-XYR1 plasmid with side-directed mutagenesis (SDM), and transforming this cassette into after PCR amplification from pGFP-XYR1$^{1-780}$ into QM9414 Δ*tku*70.

By SDM, individual nucleotides within the open reading frame (ORF) of a gene can be exchanged using pairs of mutagenesis primers containing 1-3 nucleotide mismatches. Typical features of the used oligonucleotides were: 25-40 nt length, mismatch in the middle of the primer, 50-60 % GC content, ideally no GC at the 3' end and a melting temperature of T_m > 80 °C, which was calculated using the following equation:

$$\mathbf{T_m = 81.5°C + 0.41 \times \%GC - (657/nt_{total}) - \%\ mismatch}$$

$$\text{With \% mismatch} = 100\% \times (nt_{mismatch}/nt_{total})$$

Primers (Table 3), dNTPs, 50-100 ng of template plasmid DNA, polymerase buffer and Phusion DNA polymerase (Fermentas) were combined with dH_2O in a 50µl PCR mix. The mutagenized

DNA was synthesized with 30 sec/kb in 20 repetitions in a thermal cycler, according to the recommendations of the polymerase manufacturer. 5-8µl of the 50µl PCR mix were transformed into 200µl of ultra-competent *E. coli* using a heat-shock protocol (3.1.2).

Table 3: Primer used for side-directed mutagenesis of *xyr1* locus. The mutated triplet is underlined.

Primer	5'- 3' Sequence	Nucleotides	% GC	TM [°C]
QM9136-QC-F	AGCACCCATGTGATGCTGTCCTCCACATCC	30	57	81
QM9136-QC-R	GGATGTGGAGGACAGCATCACATGGGTGCT	30	57	81

3.5.3.3 Restriction digest

The methylated template plasmid DNA was reopened by BamHI restriction digestion. The vector was digested as follows:

DNA	1µg
BamHI	3µl
1x Tango buffer	12µl

The reaction mix was filled up to a final volume of 20µl with nuclease free water and incubated for 24 hours at 37°C. After 22 hours of incubation, 3 µl of BamHI was added to the reaction mix.

3.5.3.4. Plasmid DNA amplification

Correct plasmids were amplified in *E. coli* (section 3.5.5) and used as templates for the amplification of the 9.5kb GFP-XYR1$^{1-780}$ large replacement cassettes.

3.5.3.5 PCR for amplification of the replacement cassette

The plasmid DNA was verified by control restriction digestion (see 3.5.3.3) and sequencing (Table 4). The following PCRs were carried out for amplification of pGFP-XYR1$^{1-780}$.

Reaction mix:

29 µl	PCR-H_2O
10 µl	5xPhusion HF Buffer
1 µl	dNTPs (10mM each)
0.5 µl	Phusion Polymerase (Fermentas)
1.5 µl	Primer_F (100µM)
1.5 µl	Primer_R (100µM)
1.5 µl	DMSO
5 µl	template DNA (~120 ng)
50µl	

The following PCR program was used for amplification:

3 STEP-PCR

	2 min	98°C	initial denaturation
	15 sec	98°C	melting
	30 sec	64°C	annealing
Repeat 26x	4 min	72°C	amplification (25sec/kb)
	4 min	72°C	primer extension
	∞	4°C	

Table 4: Used oligonucleotides for generation of the replacement cassette. Compatible overhangs required for recombinational cloning are indicated in bold; restriction enzyme recognition sites are underlined.

Primer Name	5'-3' Sequence
xyr1-3'-flank-F	**CCATAGTACCCTCGAG**GCAACACAACACTCACCTCTT
xyr1-3'-flank-R	**ATGCCTGCAGGTCGAC<u>GCTAGC</u>**AGAATAGGAGGATGGCTCTTG
Pxyr1-F	**CGGTACCCGGGGATCC<u>GCTAGC</u>**ACACAAGAGCAATGGCCCTAGC
Txyr1-R	**GGTAGCTCTC<u>GGGATCC</u>**ACTCGTCACACTGGCTCTCGTAC

3.5.3.6 Structure of the *gfp-xyr1*$^{\Delta 2294A}$ Replacement Cassette

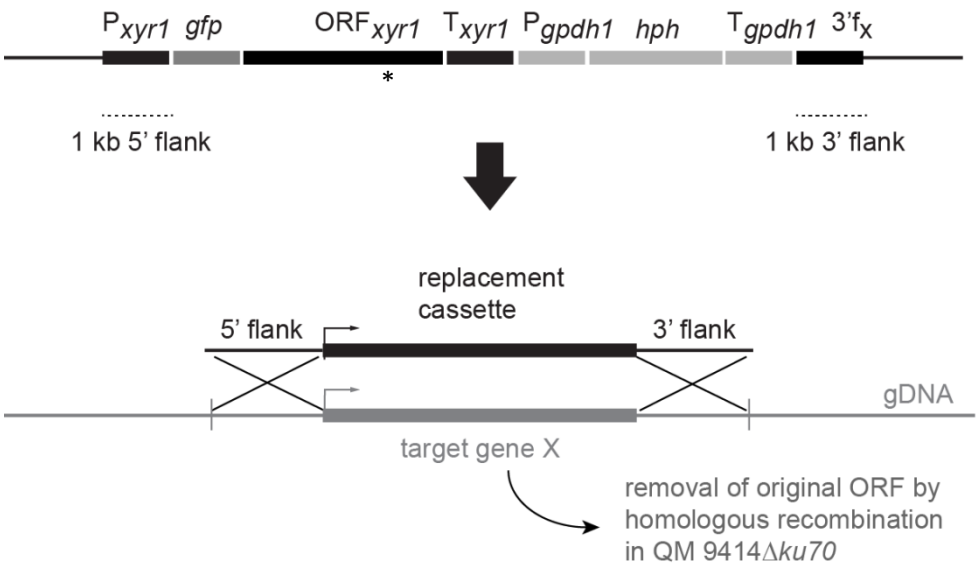


Fig. 9: Construction of the *gfp-xyr1*$^{\Delta 2294A}$ replacement cassette for transformation of QM9414. The approximate position of the Δ2294A point deletion is indicated with an asterisk.

2.4.4 DNA extraction of a preparative gel

Gel extraction was used to purify a certain band from an agarose gel. Therefore, the agarose gel was exposed to UV light and the band of interest excised using a razorblade. Gel extraction was performed following the QIAquick® Gel Extraction Kit protocol.

2.4.5 Transformation of *E.coli*

Chemically competent *E. coli* cells, stored at -80°C, were used to amplify plasmids and ligation reactions. For this, Eppendorf tubes were chilled on ice. Frozen competent *E. coli* cells were slowly thawed on ice. Thawed competent cells were gently mixed and 50µl aliquots were transferred into chilled Eppendorf tube. After addition of 5-50ng of DNA, the tubes were nudged again and immediately returned to ice for 10 minutes. Thereafter, the cells were heat-shocked for 45- 50 seconds at 42°C in a water bath. The tubes were immediately placed back on ice. 500µl of pre-warmed SOC medium was added to each transformation reaction, followed by incubation for 60 minutes at 37°C with heavy shaking. 100µl of the transformation reaction were plated on LB amp plates and incubated overnight at 37°C.

2.4.6 Minipreperation of plasmid- DNA from *E. coli*

For plasmid-DNA preparation, test tubes containing 3ml of liquid LB amp medium were inoculated with one transformed *E. coli* colony each and incubated in a rotary shaker at 37°C o/n. The next day, cells were harvested by centrifugation for 5 minutes at 13000 x g at room temperature.

After removal of the supernatant, 100µl of buffer P1 were added to the cells, followed by mixing on a Vortex shaker. Upon that, 200µl of buffer P2 were added to the tubes, which were subsequently inverted for 5-8 times, and allowed to incubate for not more than 5 minutes at room temperature. To stop lysis, 200µl of buffer P3 were added and the tubes were inverted again for 5-8 times. After centrifugation for 12 minutes at 13000 x g at room temperature, the supernatant was cleared in a fresh Eppendorf tube prepared with 420µl Isopropanol and incubated at -20°C for 20 minutes. Upon centrifugation for 30 minutes at 13000 x g and 4°C or room temperature, isopropanol was replaced with 70% ethanol to wash the DNA pellet. After a final centrifugation step, the ethanol was discarded, the DNA pellet air-dried for 5 minutes and resuspended in 30-50ml sterile water. DNA concentration was determined using the 'Nanodrop' device.

Buffer P1:

50mM Glucose	9.91 g/l
10mM EDTA	3.72 g/l
25mM Tris	3.3 g/l

pH 8,0 with HCl

Buffer P2:

0,2M NaOH	8g/l
SDS	1% (w/v)

Buffer P3:

3M Potassium acetate	294.32 g/l
2M Pure Acetic acid	11.5% (v/v)

pH 4.8

2.4.7 Midipreparation of plasmid-DNA from *E. coli*

In order to upscale the yield of pDNA, plasmids were purified from *E. coli* using the Pure YieldTM Plasmid Midiprep kit (Promega), following the protocol of the manufacturer.

2.4.8 Purification of DNA by isopropanol precipitation

In order to precipitate DNA, Isopropanol (2x of the sample volume at hand) and 3M NaOAc pH= 5.2 (0.1x of the sample volume at hand) were added to the respective Eppendorf tube. After that, an incubation for 20min at -20°C and centrifugation for 30min at 10000 x (Hickey) and 4°C followed. The supernatant was discarded. 300µl of 70% ethanol were added to wash the pellet. Then a centrifugation at 10000 x g for 5 minutes followed. The supernatant was discarded again and the pellet was dried. Upon that, the pellet was resolved in 30-50µl of nuclease free water or EB buffer.

2.4.9 Transformation of *T. reesei* protoplasts

A freshly sporulated plate is essential for protoplast preparation. After adding 15ml of NaCl/ Tween solution to the plate and displacing the conidia with a Drigalski spatula, the spore solution was filtered through a tube filled with glass- wool. Prior to that, PDA plates with autoclaved cellophane discs were prepared. Per PDA plate, one cellophane disk was put on the solid medium.

As a rule of thumb, 10 cellophane- covered PDA plates are needed for one round of transformation. Then, 60µl of the filtered spore solution were pipetted on the cellophane disk, streaked out and incubated overnight at 28°C. The next day, the cellophane was covered with freshly grown mycelium. So that, 0.45g of lysing enzyme (*Trichoderma harzianum*, Sigma #L-

1412) were added to 50ml of solution A and mixed gorgeously. After sterile filtration through a 0.2µm filter, 25ml of this solution were transferred to a sterile petri dish. The mycelium was covered with cellophane disks and was put into the petri dish containing the lysing enzyme solution. Then the petri dishes were incubated for 10 minutes.

After that, the mycelium could be separated from the cellophane by using a forceps. Cellophane disks were discarded. The petri dish containing the mycelium was incubated at 28°C for approximately 90-120 minutes to generate protoplasts. Afterwards the protoplast solution was filtered through a glass- wool- funnel and washed with solution A. This step was followed by 10 minutes of centrifugation at 644 x g at 4°C with the swing- out- rotor.

From this point on, the protoplast solution was kept on ice. The supernatant was discarded and the pellet was dissolved in 4ml of 4°C solution B. Another round of centrifugation was carried out. The supernatant was discarded again and the pellet was dissolved in 200µl of 4°C solution B. For dissolution, the tube was nudged gently to prevent protoplast bursting. Approximately 30µg of precipitated DNA and 50µl PEG were added the protoplast solution. The tube was nudged gently and incubated on ice for 20 minutes. Subsequently, 2ml PEG were added and gently mixed with the solution. After incubation for 5 minutes at room temperature, 4ml of room tempered solution B were added. 4ml of overlay medium were transferred into a test tube and tempered in a 48°C water bath. 300µl of the protoplast-containing solution were added to the overlay medium.

The liquids were mixed by rolling and poured on plates containing the bottom medium. At last, the plates were incubated at 28°C and checked for transformants during the following days.

2.4.10 Quick extraction of DNA from fungal transformants

For the determination whether obtained transformants contained the respective transforming DNA, a DNA quick extraction was carried out. Therefore, a small piece of mycelium was mechanically disrupted by means of glassbeads and thorough vortexing. After incubating for 15 minutes and shaking at 200rpm at 65°C, 150µl of buffer P1 was added, followed by vortexing. Upon centrifugation for 1 minute at 10000 x g, the supernatant was transferred into a new Eppendorf tube containing 500µl of isopropanol. After another round of vortexing and centrifugation for 10 minutes at 10000 x, the isopropanol was removed. The DNA pellet was washed with 75% ethanol. After that, the pellet was dried and resuspended in 30µl of nuclease free water.

Lysis Buffer:

50mM glucose	9.91 g/l
10mM EDTA	3.72 g/l
25mM Tris	3.3 g/l
pH 8,0 with HCl	

2.4.11 PCR-genotyping of *T.reesei* transformants

To confirm that selected *T. reesei* transformants have integrated the respective transformation cassettes, e.g. for the expression of TRAL002 fusion constructs, at the intended locus, a PCR-based genotyping procedure was used as explained in detail elsewhere (Lichius et al. 2012). In order to check integration of the transforming DNA in the obtained transformants, genomic DNA was analyzed using a PCR- based approach.

Firstly, primers were chosen accordingly to be able to check if all three parts of the replacement cassette (*gfp*, *xyr1Δ2294A* and *hph*) integrated successfully at the intended locus (Table 5). If the result was positive, single spore isolation was carried out following a PCR with genotyping primers. These primers amplify the whole replacement cassette and are a true indicator whether the gene was replaced successful at the correct locus**.**

Table 5: Oligonucleotides used for PCR-genotyping.

Primer Name	5'-3' sequence
5Pxyr1-ver-F	CCAGCTGCCACTCTCATG
Txyr1-ver-R	AGTCGCTCATGATCCTACCAG
Pxyr1-ver-F	AAGGATGCCGACTTAACGAAC
Xyr1-ver-R	CCTGGCAGCAATAAGAGAGC

PCR reaction mix:

15 µl	PCR-H_2O
5 µl	5xPhire HF Buffer
0.5 µl	dNTPs (10mM each)
0.5 µl	Phire Polymerase (Fermentas)
1.5 µl	Primer_F (100µM)
1.5 µl	Primer_R (100µM)
1 µl	genomic DNA (~80ng)
25 µl	

The following PCR program was used for amplification:

2STEP-PCR		2 min	98°C	initial denaturation
		15 sec	98°C	melting
	repeat 29x	4 min	72°C	amplification(25sec/kb)
		4 min	72°C	primer extension
		∞	16°C	

2.4.12 Preparation of labware and biomass for subsequent RNA isolation

RNA isolation demands overnight DEPC (diethylpyrocarbonat solution 0.1%) treatment of all materials needed, subsequent autoclaving and drying at 120°C.

Mycelium derived from replacement experiments in submerged cultures was harvested, thereafter immediately frozen in liquid nitrogen and stored at -80°C.

2.4.13 Extraction of RNA and cDNA synthesis

All steps were performed on ice or in liquid nitrogen as indicated:

In 2mL Eppendorf tubes a mix of 700µl Chirgwin-reagent and 6µl mercaptoethanol was prepared. Then a small amount of frozen mycelia was ground with pestle and mortar in liquid nitrogen. Biomass powder was added to the 2 ml tube with Chirgwin/mercaptoethanol and was vortexed for 10-15 sec until a homogenous emulsion has formed. Stepwise 70µl 2M NaAc (pH 4.0), 700µl Phenol (pH 4), 400µl chloroform- isoamylalcohol mix was added and vortexed well after every addition. After that, the mixture was incubated on ice for at least 15–20 min. The tubes were centrifuged for 20 min at 14.000 g and 4°C

From here work was continued in an RNAse free environment (DEPC treated tips, tubes, water):

Eppendorf tubes with 700µl Isopropanol were prepared on ice. The aqueous phase (upper phase; max. 700 µl) was transferred to the isopropanol containing tube and the tube must be inverted a few times to mix gently. RNA was precipitated for at least 1h at -20°C. After that, RNA was formed to a pellet by centrifugation for 20 min at 14.000 g and 4°C.Then the supernatant was discarded and the pellet was washed twice with 1 mL pre-cooled 70% EtOH to remove as much EtOH as possible. Then air dry pellet at RT for 5-10 min (do not overdry). The pellet was dissolved in 50µl DEPC water (dissolution is facilitated by incubation for 10 min at 55-60°C in a thermo block.) Then RNA concentration (NanoDrop) could be determined. If RNA was not immediately processed, it could be stored at -80°C.

DNaseI digested RNA samples were reverse transcribed into cDNA by means of RevertAidTM H Minus First Strand cDNA synthesis kit (Fermentas). The procedure was carried out following the manufacturer's manual. Instead of using either oligo(dT)18 primer or random hexamer primer, a combination of both primers was used. Incubation steps were carried out regarding the random hexamer primer.

2.4.14 RT- qPCR

In order to measure the amount of specific gene transcripts, quantitative PCR was used. Relative transcript levels were determined by normalizing the absolute expression levels of *cbh1*, *cre1*, *xyn2*, *xyr1* marker genes onto the house-keeping reference gene translation elongation factor 1 (*tef1*) of *T. reesei*. Reactions were performed in triplicates and at least two biological replicates using a BIO-RAD® iQ™ thermal cycler 1.

PCR reaction mix:

SYBR green mix	12.5 µl
PCR-H2O	8.5 µl
primer F + R (1:16)	2 µl
cDNA template (1:100)	2 µl
	25 µl

The following PCR program was used for amplification:

1 cycle:	95°C	2min
50 cycles:	95°C	20sec
	60°C	20sec
	72°C	20sec

Reaction mixes, as well as no-template controls (to control for unspecific background products) were prepared for each primer pair (Table 6). All samples were prepared as triplicates. 23µl of the respective reaction mixes were filled into the wells of a 96- well plate (Eppendorf). Thereupon, 2µl of the corresponding cDNAs were added. The plates were sealed with optical foils, shortly centrifuged to spin down the liquids and placed in the

Table 6: Oligonucleotides used for RT-qPCR.

Target gene	Primer Name	5' - 3' Sequence	Efficiency
translation elongation factor 1 *tef1*	qPCR-tef1-F	CCACATTGCCTGCAAGTTCGC	95%
	qPCR-tef1-R	GTCGGTGAAAGCCTCAACGCA	
xylanase 2 (*xyn2*)	qPCR-xyn2-F	CAACCAGCCGTCCATCATCG	97%
	qPCR-xyn2-R	ATCGTCCCGAGCGTCAGG	
carbon catabolite repressor 1 (*cre1*)	qPCR-cre1-F	GTCTGAGAAACCTGTCCCTG	91%
	qPCR-cre1-R	GGCTAATGATGTCGGTAAGTG	
cellobiohydrolase 1 (*cbh1*)	qPCR-cbh1-F	CCGAGCTTGGTAGTTACTCTG	98%
	qPCR-cbh1-R	GGTAGCCTTCTTGAACTGAGT	
xylanase regulator 1 (*xyr1*)	qPCR-xyr1-F	CCATCAACCTTCTAGACGAC	100%
	qPCR-xyr1-R	AACCCTGCAGGAGATAGAC	

2.5 Sequencing

In order to check *in locus* integration of the transforming DNA in the obtained transformants, DNA amplicons of that region (see 3.4.11 for details) were sent to sequencing. Two primers were used specifically for sequencing (Table 7).

Table 7: Primer used for DNA sequencing. Pxyr1-seq-fw binds 150bp upstream of promoter end and xyr1-seq-fw in the center of xyr-1 open reading frame.

Name	5'-3' sequence	Nucleotides	% GC	Tm [°C]
Pxyr1-seq-fw	GACAGCAGCAGTAGTCAGGT	20	55	60
xyr1-seq-fw	CCGTCTCCCAAGACTAGC	18	61	60

2.6 Microscopy

2.6.1 Sample preparation for live- cell imaging

To quantify GFP-labeled transcription factors expression and subcellular localization must be determined. Therefore a scanning confocal microscopy and image analysis was done. Fungal cells were prepared as germling samples from submerged flask cultures (see 3.4.4). In the next step 20µl cell suspension was placed between two glass cover slips or as mycelial samples from growing plate (see 3.4.3) using the 'inverted agar block' method. (Hickey et al. 2005)

2.6.2 Confocal microscopy

Live cell imaging was performed using a Nikon C1 confocal laser scanning unit mounted on a Nikon Eclipse TE2000-E inverted microscope base (Nikon GmbH, Vienna, Austria). GFP-labelled proteins were excited with the 488nm line of an argon ion laser, and the emitted fluorescence light separated by a Nikon MHX40500b/C100332 filter cube and detected with a photo-multiplier tube in the range of 500-530nm. Laser intensity and laser dwell time during image recording were kept to a minimum to reduce photobleaching of GFP-labelled proteins and phototoxic effects. A Nikon Plan Apo VC 60x/1.2 water immersion objective lens was used for microscopy. With a Nikon C1-TD transmitted light detector mounted behind the condenser turret, bright field pictures were captured simultaneously. Images were recorded with a maximum resolution of 1024x1024 pixels and saved as TIFF. The ImageJ software platform (http://rsb.info.nih.gov/ij/) was used for quantitative image analysis and figure preparation. Apart from display range adjustments and cropping, images were not subjected to further manipulation. Absolute fluorescence intensities were displayed and statistically evaluated using Microsoft Excel. Interactive surface plots were done with the corresponding plugins of MacBiophotonics ImageJ work package available at http://www.macbiophotonics.ca/software.htm.

2.6.3 Quantitative image analysis

T. reesei is a multinucleate organism and its nuclei are not mitotically or transcriptionally synchronized. To account for the high degree of variation in gene regulation between individual nuclei within the nuclear population of imaged hyphae, the nucleo-cytoplasmic fluorescence ratio (n/c-ratio) was determined for each tested condition. The ration between the average nuclear fluorescence intensity within the population of nuclei and the average cytoplasmic fluorescence intensity between these nuclei is represented by this value, and

indicates how much stronger the average fluorescence signal is in nuclei compared to their surrounding cytoplasm. Therefore, changes in the n/c-ratio instead of comparing absolute changes accounts for the local differences in transcription factor expression and subcellular localization within the imaged hyphal population were compared. The residual auto-fluorescence background and the detection limit of the microscope were also taken into account.

3 Results

3.1 The truncated form of XYR1 present in QM9136 is non-functional and responsible for the cellulase negative phenotype.

3.1.1 Generation and analysis of different *xyr1* recombinant strains for the analysis of the role of XYR$^{1-780}$ in cellulase regulation.

Analysis of the genome sequence of QM9136 revealed that the cellulase and xylanase regulator XYR1 is present in a truncated termed XYR$^{1-780}$ which could be responsible for its cellulase negative phenotype. To investigate the role of the truncated form of XYR1 present in QM9136 in cellulase expression, a number of different strains was constructed or analyzed. These recombinant strains were (i) a QM9136 strain expressing a native *xyr1* to analyze if the *xyr1* mutation in QM9136 is the only mutation responsible for the cellulase negative phenotype, (ii) a Δ*xyr1* strain expressing the truncated form of XYR$^{1-780}$ to see if this truncation has any influence in the absence of the native XYR1 and (iii) a GFP-XYR$^{1-780}$ expressing strains to visualize the intracellular location of the truncated form of XYR1.

Plasmid pGFP-XYR1$^{1-780}$ expressing the truncated form of XYR1 was generated from pGFP-XYR1 by side-directed mutagenesis using primers QM9136-QC-F and QM9136-QC-R to delete the adenosine at position 2294 in the *xyr1* gene. The point deletion was verified by DNA sequencing. The corresponding gene replacement cassette was amplified from pGFP-XYR1$^{1-780}$ and transformed into QM9414 *Δtku*70. Strains where the cassette had integrated at the endogenous *xyr1* locus thereby exchanging *xyr1* by the mutated allele were named TRAL007 1-3 (Fig. 10).

In addition to this newly generated strain expressing the truncated GFP-XYR1$^{1-780}$ construct in a QM9414 background, two previously prepared types of transformants including QM91.12, QM91.20 and QM91.27 expressing XYR1 from QM9414 in a QM9136 background, and XYR26 and XYR27 expressing XYR1 from QM9136 in a Δ*xyr1* background, were analyzed. Their genetic and functional verification included PCR-genotyping, DNA-sequencing of the amplicon, growth tests on various carbon sources and an expression analysis of specific key genes. These methods were intended to illustrate the functionality or loss-of-function of the XYR1 or XYR1$^{1-780}$ in the different strain backgrounds.

3.1.2 Genotyping and DNA sequencing

Integration of the transformation cassette within or outside the *xyr1* locus was analyzed by PCR-genotyping and further verified by DNA sequencing of the resulting amplicons.

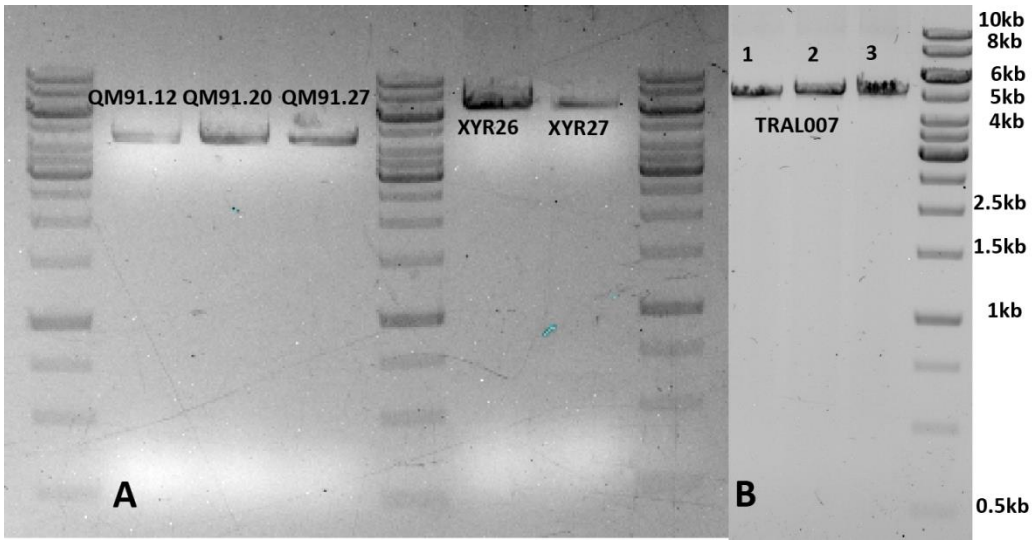


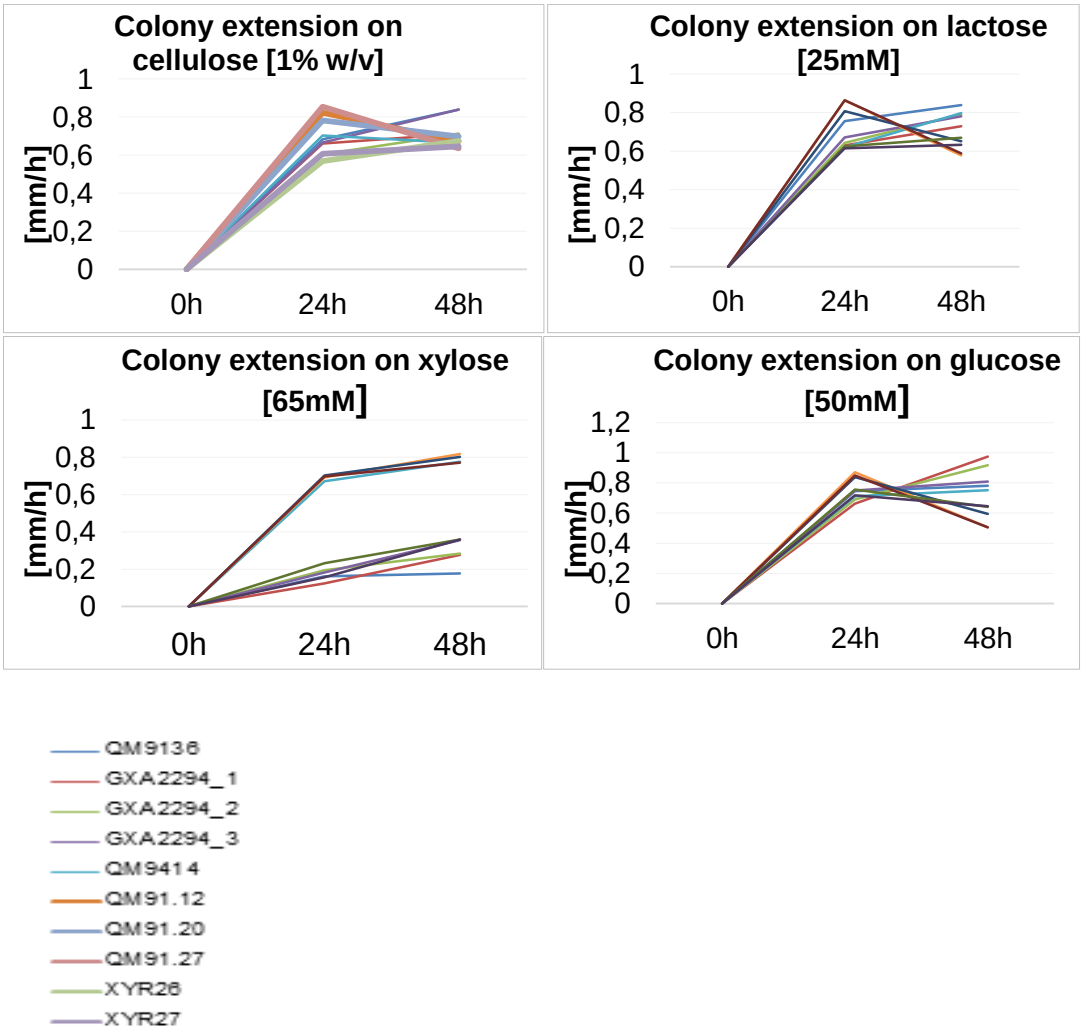
Fig. 10: Gel electrophoresis analyses of *xyr1* locus-specific PCR products confirming the integration of different alleles of *xyr1* in *T. reesei* transformants. (A) Strains QM91.12, QM91.20, QM91.27 (QM9136 transformed with *xyr1* locus from QM9414) and strains, XYR26 and XYR27 (Δ*xyr1* transformed with mutated *xyr1* from QM9136). **(B)** Strains TRAL007 1-3. Used primers were 5Pxyr1-ver-F and Txyr1-ver-R. A 4.5 kb amplicon corresponds to the native *xyr1*, a 5.2 kb fragment to a *gfp-xyr1* fusion and a 6 kb amplicon to a *xyr1* deletion by *amdS* present in the Δ*xyr1* strain.

Genomic DNA isolated from each clone was used as template for the amplification of specific fragments spanning the whole *xyr1* promoter to terminator region (Fig. 10). In case a *gfp-xyr1* fragment has been integrated upon replacement of the native locus, a larger amplicon (5.2 kb) was to be expected compared to the native locus (4.5 kb). Only for strains TRAL007 1-3 this 5.2 kb band was detected, whereas from QM91.12, QM91.20 and QM91.27 a 4.2 kb band was detected as the native and mutated *xyr1* locus produce the same fragment. In the XYR26 and XYR27 strains an about 6 kb amplicon could be recovered resulting from the replacement of the *xyr1* coding region by the *amdS* gene (Stricker et al. 2006). This suggested that in TRAL007 strains gene replacement occurred at the endogenous locus, whereas in the two other strains transformation an ectopic integration of *xyr1* was observed. These findings were additionally verified by DNA sequencing of the purified amplicons. For TRAL007, the intended A2294 deletion in the *xyr1* locus was obtained, whereas in the QM91.12, QM91.20 and QM91.27 strains both the native and mutated *xyr1* sequence was present. In strains XYR26 and XYR27 only the mutated form of the XYR1 was present.

3.1.3 Growth test on 65mM xylose medium

Colony extension growth tests on different carbon sources were conducted in order to identify XYR1 loss-of-function mutants which grow poorly on high concentrations of xylose (e.g. 65mM) (Fig. 11 A and B). The transformant strains TRAL007 1 -3, XYR26, XYR27, QM91.12, 20, 27, as well as the reference strains QM9136 and QM9414 were compared. Strong growth retardation defects on 65mM xylose were found for strains QM9136, all TRAL007 strains, XYR26 and XYR27, and QM91.27. QM91.12 and QM91.20, on the other hand, were indistinguishable from the reference strain QM9414, confirming that the mutated *xyr1* allele in QM9136 has successfully been complemented by the additional, ectopically integrated, functional copy of wild-type *xyr1* (Fig. 11).

A



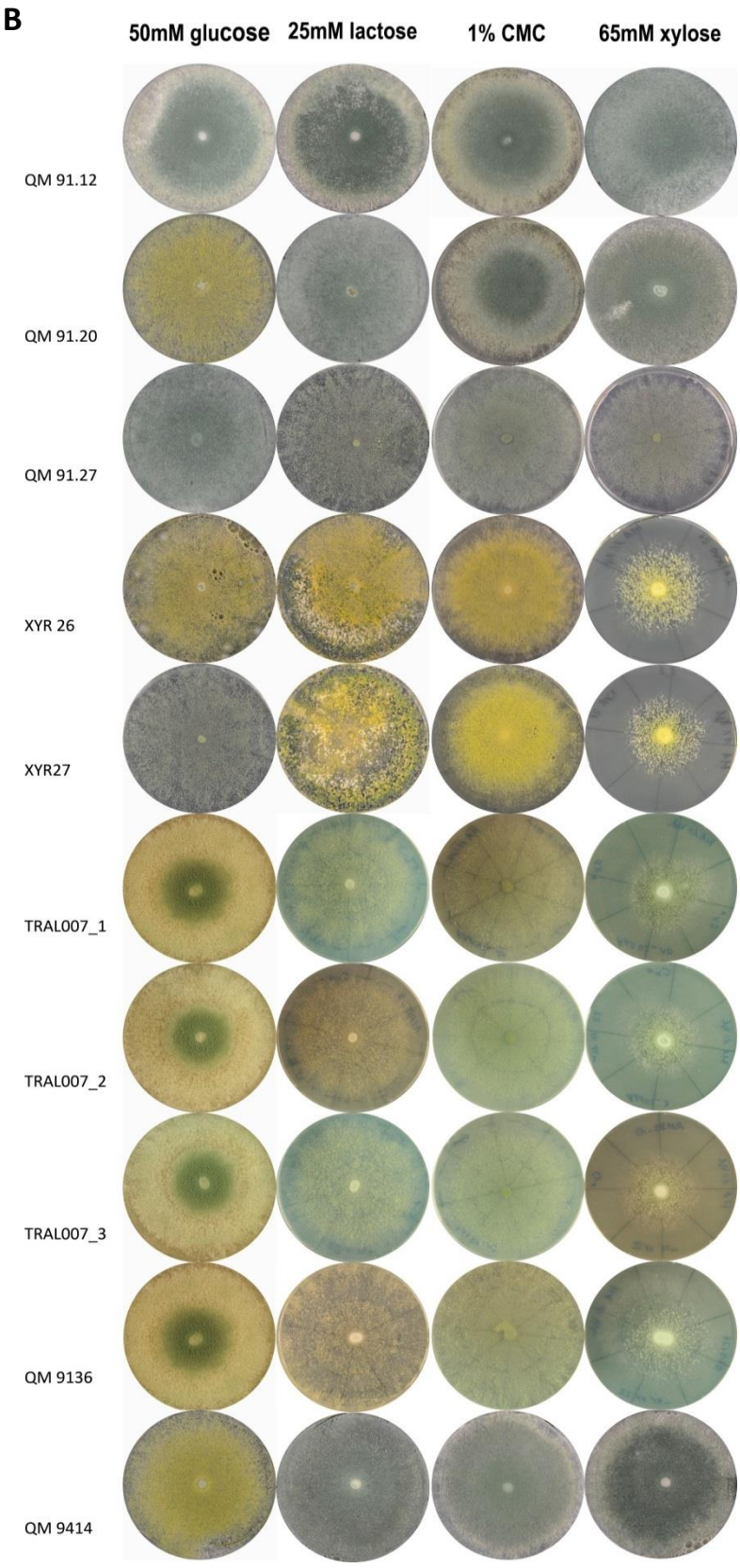


Fig. 11: Colony phenotypes of xyr1 recombinant strains and reference strains. (A) Colony extension rates on various carbon sources. Notable is the growth inhibition of XYR1 loss-of-function strains on 65mM xylose. Limitations of the plate diameter are responsible for the decreases of colony extension in QM91.27 and QM91.20 after 24h. **(B)** Macroscopic colony phenotypes of strains measured in (A). Pictures were taken after 72 h incubation at 28°C with 12/12 light/dark cycles.

This data also confirmed that the truncated XYR1 of QM9136 is incapable of complementing the deletion of *xyr1* in the Δ*xyr1* strain, and that, furthermore, QM91.27 still has the parental QM9136 phenotype, and thus the reintroduction did not result in complementation of the Δ*xyr1* phenotype. On 50mM glucose the growth phenotypes were similar while on 25mM lactose and 1% cellulose colony extension was similar between strains although biomass formation is reduced on plates as seen by the thickness of the mycelium or when these strains are grown in liquid medium (unpublished data).

3.1.4 Gene transcription analysis

The expression of different key genes for cellulase production (*cbh1*, cellobiohydrolase 1), xylanase production (*xyn2*, xylanase 2), and general transcriptional activity associated to either induction (*xyr1*) or repression (*cre1*) were analyzed using RT-qPCR. Inducing conditions were generated through carbon source replacement to 1.4mM sophorose (Fig. 12).

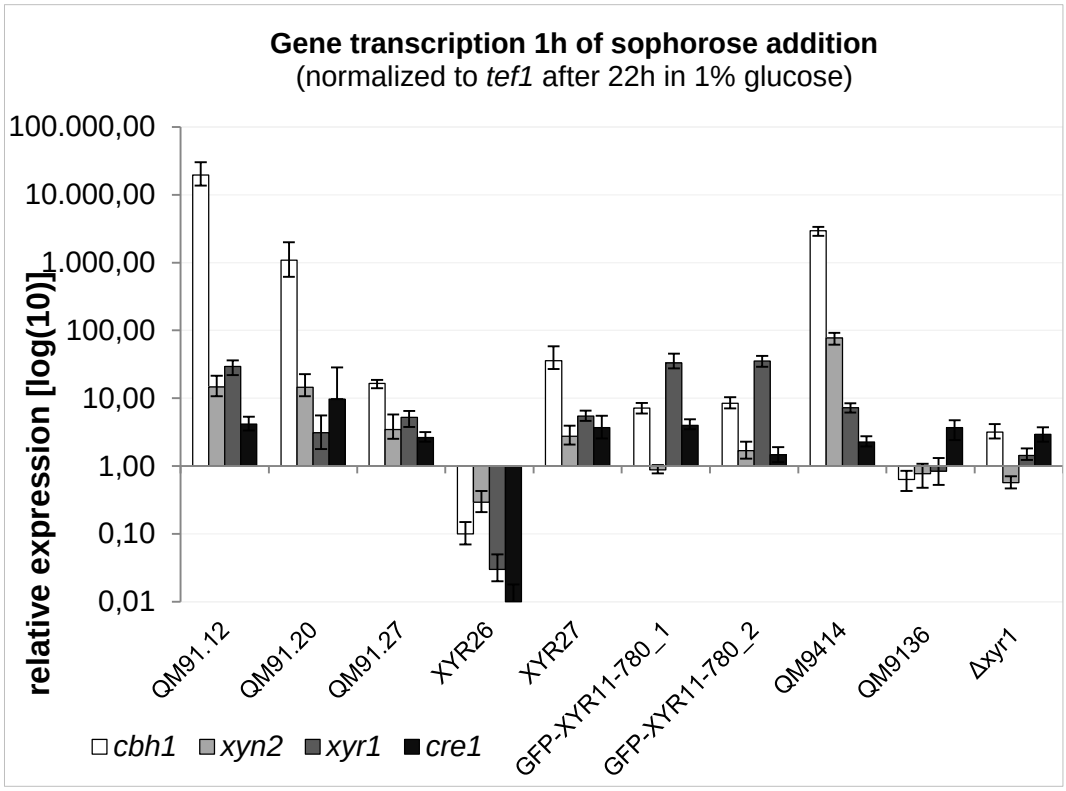


Figure 12: *Cbh1*, *xyn2*, *xyr1* and *cre1* transcription analysis in different *xyr1* recombinant strains. Strains QM91.12, QM91.20, QM91.27 (QM9136 + *xyr1*), XYR26, XYR27 (Δ*xyr*1 + *xyr1*$^{\Delta 2294A}$), TRAL007 strains 1-2 (QM9414 + *gfp-xyr1*$^{\Delta 2294A}$) and the reference strains QM9136, QM9414 and Δ*xyr1* were grown for 22 h on glucose before their mycelium was replaced to 1.4mM sophorose. All expression data are normalized to *tef1* expression after 22h in 1% glucose.

After 1h of induction with 1.4mM sophorose (Fig. 12), *cbh1* was found to be upregulated in QM91.12, QM91.20 and in the reference strain QM9414, and correlated with upregulated *xyr1* expression in QM91.12 and QM9414. Together, it indicates a successful XYR1-mediated cellulase induction with sophorose. All other strains showed only very low or downregulated *cbh1* and *xyr1* expression level, suggesting that sophorose induction had no significant effect on cellulase expression in strains where only a mutated copy of *xyr1* is present. *Cre1* was generally expressed at a very low level, with relative expression never reaching higher than 10-fold in all strains which is in line with cellulase-inducing conditions. Notably, QM91.27 and XYR27 showed a gene transcription pattern distinct from those of the other strains or parental strains. This suggested that the *xyr1* is not present anymore or inactive, or that the integrated cassette inserted in a locus where its expression was atypically regulated. In summary, this analysis suggested that only QM91.12, QM91.20, XYR26, TRAL007-1, TRAL007-2 are correctly engineered transformants which respond in the expected way, and that strains QM91.27 and XYR27 will not be useful for further studies. The most important output from this analysis, nevertheless, is that genetic rescue of QM9136 with the *xyr1* wild-type copy successfully restored XYR1-mediated cellulase expression in the two strains QM91.12 and QM91.20, whereas expression of the truncated XYR1 alleles in a *xyr1* deletion background in XYR26 or the truncated GFP-XYR1 in strains TRAL007 1-2 does not induce cellulase expression.

In summary, the applied transformant verification strategy confirmed that: (1) in QM91.12 and QM91.20 the wild type *xyr1*, although ectopically integrated, successfully rescued the XYR1-loss-of-function phenotype of QM9136, (2) the ectopic expression of the truncated *xyr1* allele form QM9136 in XYR26 does not alter the mutant phenotype of Δ*xyr1*, (3) endogenous integration of the engineered *gfp-xyr1*$^{\Delta 2294A}$ construct in a QM9414 background results in a XYR1-loss-of-function phenotype identical to that of QM9136. It is important to note, that strains QM91.12 and QM91.20, due to the ectopic integration of the wild type *xyr1* allele, are heterogenic for *xyr1*, i.e. truncated and full-length *xyr1* are co-expressed. This might explain the slightly altered gene expression profiles of these strains in comparison to the QM9414 reference strain.

3.2 XYR1 shuttling dynamics and correlated gene transcription

In order to observe the intracellular shuttling dynamics of transcription factors by means of live-cell imaging, tagging with fluorescent proteins, such as the green fluorescent protein (GFP), is the method of choice. Functionality of a GFP-XYR1 construct in *T. reesei* QM9414 has recently been demonstrated (Lichius et al. 2014) and in this study the shuttling dynamics of the truncated XYR^{1-780} was established.

3.2.1 Shuttling of the GFP-XYR1 is not influenced in QM9136 compared to QM9414.

In the first experiment the role of the genetic background of QM9136 on the shuttling dynamics of a full length GFP-XYR1 fusion was investigated. Strains expressing GFP-XYR1 in a QM9136 background were available and were compared to strains where the same construct was introduced in QM9414. In both strains GFP-XYR1 was expressed form the *xyr1* locus.

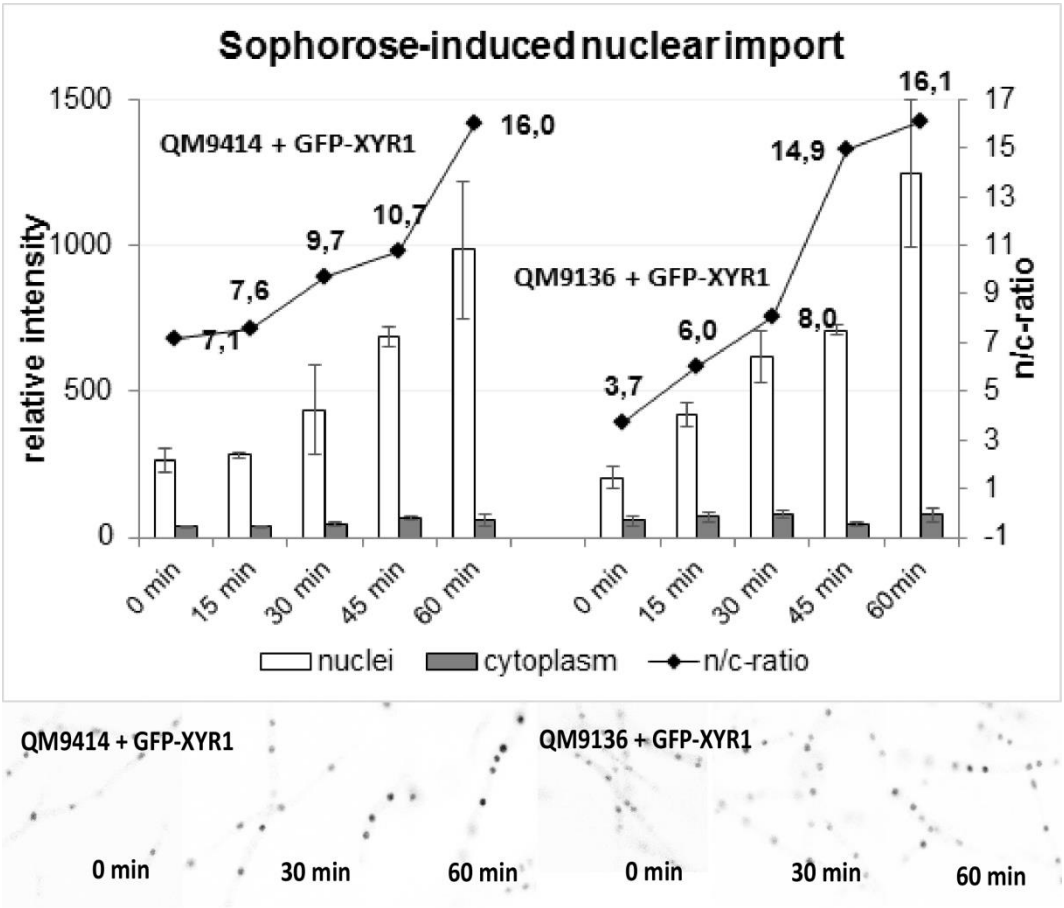


Fig. 13: Sophorose-induced nuclear import of GFP-XYR1 in QM9414 and QM9136 transformants. Nuclear import of GFP-XYR1 after replacement of a glucose culture to sophorose was followed for 60 min. Fluorescence intensity was measured in the cytoplasm and the nuclei. Representative, inverted fluorescence images of GFP-XYR1 accumulation in nuclei at the indicated time points of the experiment are shown below the graph.

Upon sophorose induction, XYR1 shuttles equally well and with similar kinetics into nuclei of both QM9414 and QM9136 (Fig. 13). In addition to this the transcription of different key genes indicative for cellulase expression was studied. As seen in Fig. 14 the resulting gene expression profiles are comparable between both GFP-XYR1 expressing strain of the QM9414 and QM9136 strain line. A comparison of the gene expression profile to the reference strains also shows that the results are in line with those of the QM9414 strain, however, distinct from those of the XYR1-defective QM9136 and Δ*xyr1* negative strains (Fig. 14).

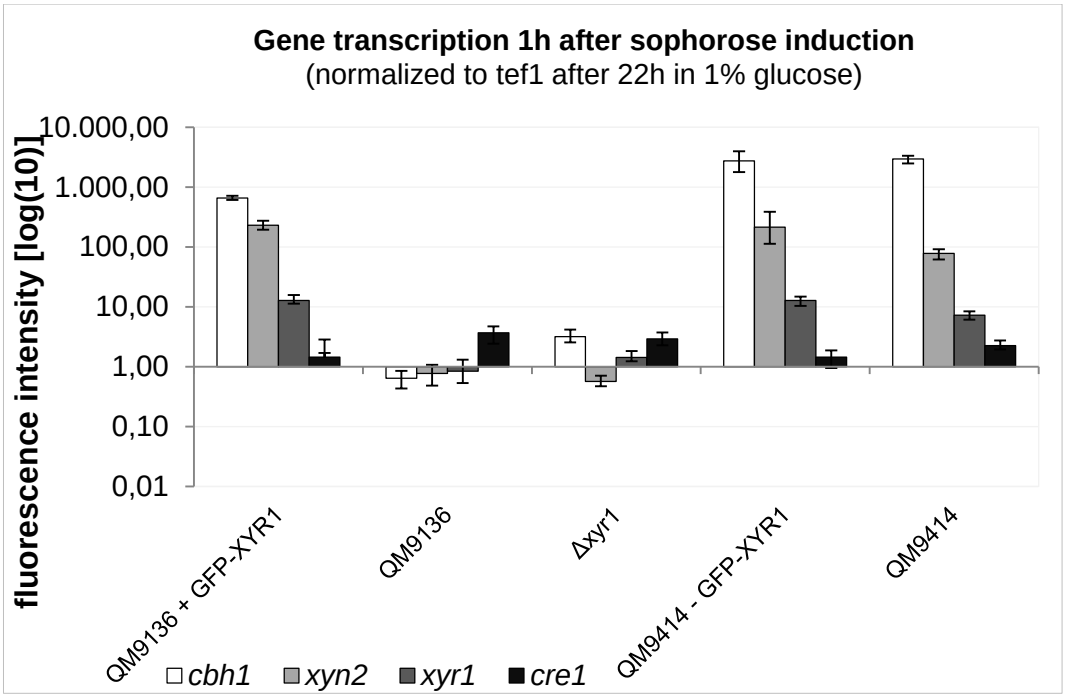


Fig. 14: Gene transcription after 1h of 1.4 mM sophorose addition. Expression of GFP-XYR1 rescues the cellulase-negative phenotype of QM9136, and functions as the wild type XYR1 in a QM9414 background. Δ*xyr1* is the XYR1-negative control strain, showing a similar gene expression profile as QM9136.

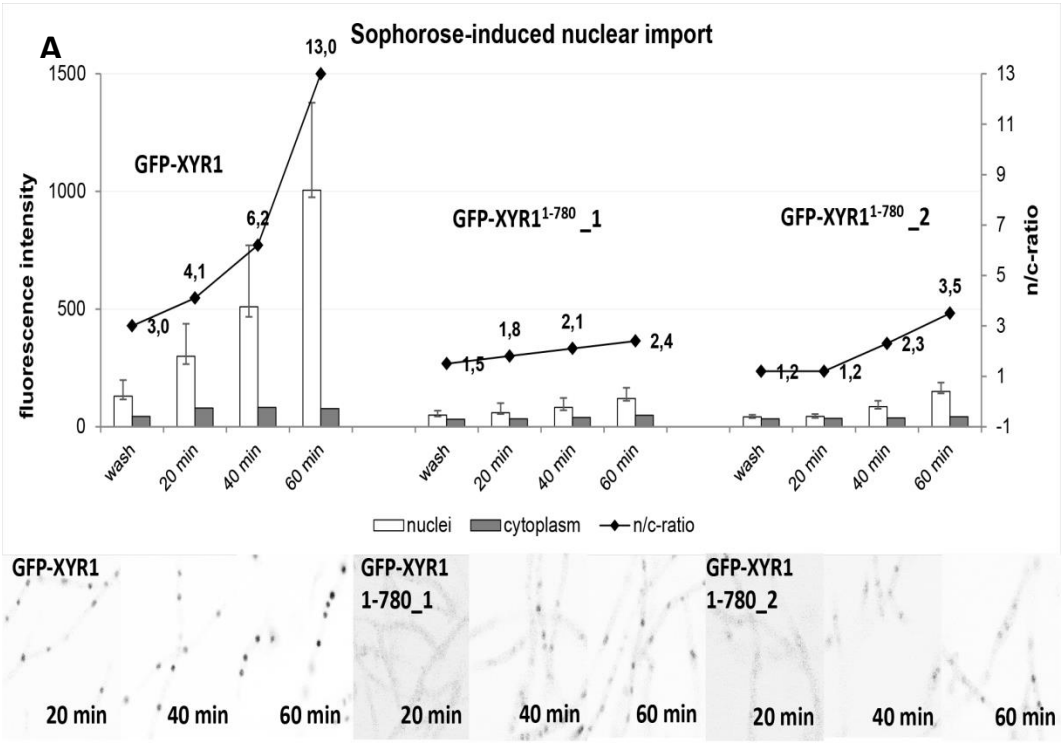
As expected, *cbh1* expression in QM9414 + GFP-XYR1, QM9136 + GFP-XYR1 and QM9414 Δ*tku70* is up-regulated following induction by sophorose. This is in accordance with the nuclear import of XYR1 which is clearly recognizable after 60 min of sophorose addition. *Cbh1* and *xyr1* expression in QM9136, on the other hand side, is on a low level compared to the resulting from the point mutation in *xyr1*. Similar results are found for the Δ*xyr*1 strain because of the complete absence of *xyr1*.

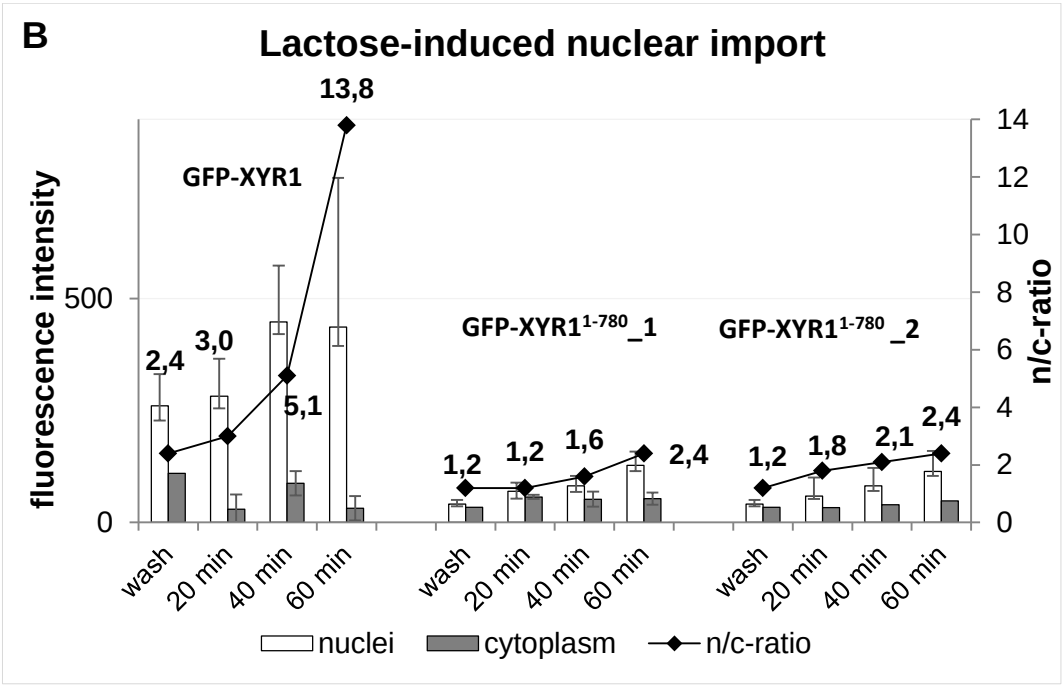
These data confirm that QM9136 strain background has no influence on the function of the GFP-XYR1 fusion constructs with respect to nuclear import dynamics and that the GFP-XYR1 is able to stimulate transcriptional activity of the *cbh1* gene. In addition, this data verify the

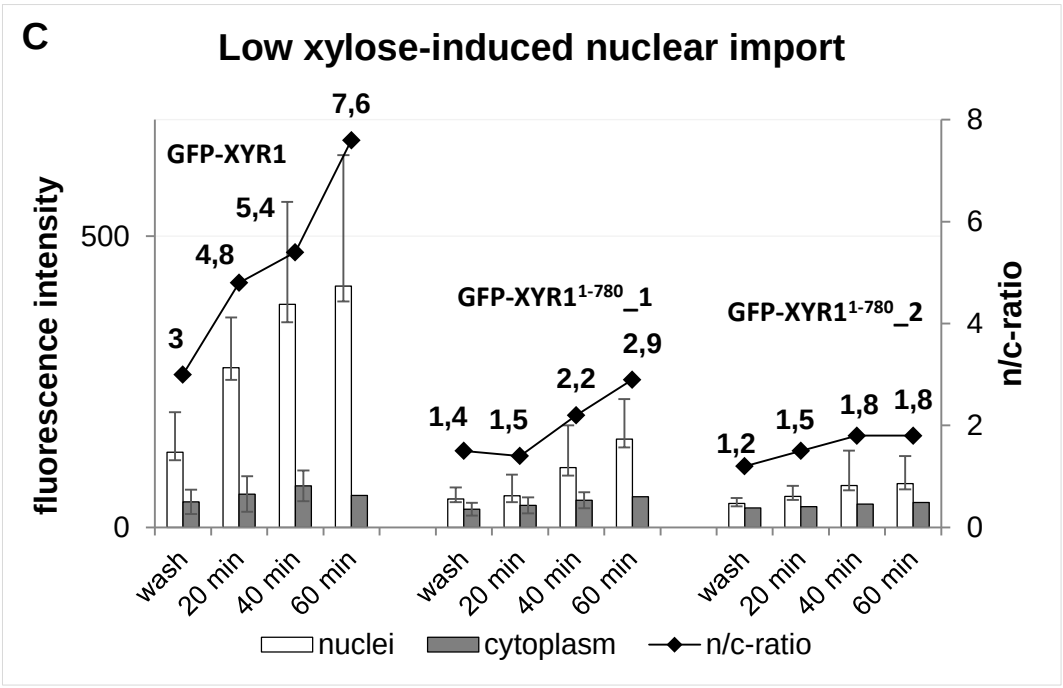
earlier observation, that the genetic replacement of the mutated *xyr1*$^{\Delta 2294A}$ allele by a native *xyr1* can fully compensate the *xyr1* loss-of-function phenotype of QM9136.

3.2.2 Influence of XYR1 truncation on shuttling dynamics

Next, the influence of the C-terminal truncation of XYR1 on the shuttling dynamics was studied to see whether the putative regulatory domain at XYR1's C-terminus is involved in the nuclear import mechanism itself. Therefore, we compared the shuttling dynamics of GFP-XYR1$^{1-780}$ with that of the native GFP-XYR1 (Fig. 15) in a QM9414 strain background. All strains were grown as submerged pre-cultures on 50mM glucose as carbon source, and were then replaced to 1.4mM sophorose, 25mM lactose, 1mM and 65mM xylose.




B
Lactose-induced nuclear import
fluorescence intensity
n/c-ratio
GFP-XYR1
GFP-XYR1$^{1-780}$_1
GFP-XYR1$^{1-780}$_2
13,8
5,1
3,0
2,4
1,2
1,2
1,6
2,4
1,2
1,8
2,1
2,4
wash
20 min
40 min
60 min
nuclei
cytoplasm
n/c-ratio


C
Low xylose-induced nuclear import
fluorescence intensity
n/c-ratio
GFP-XYR1
GFP-XYR1$^{1-780}$_1
GFP-XYR1$^{1-780}$_2
7,6
5,4
4,8
3
2,9
2,2
1,4
1,5
1,2
1,5
1,8
1,8
wash
20 min
40 min
60 min
nuclei
cytoplasm
n/c-ratio

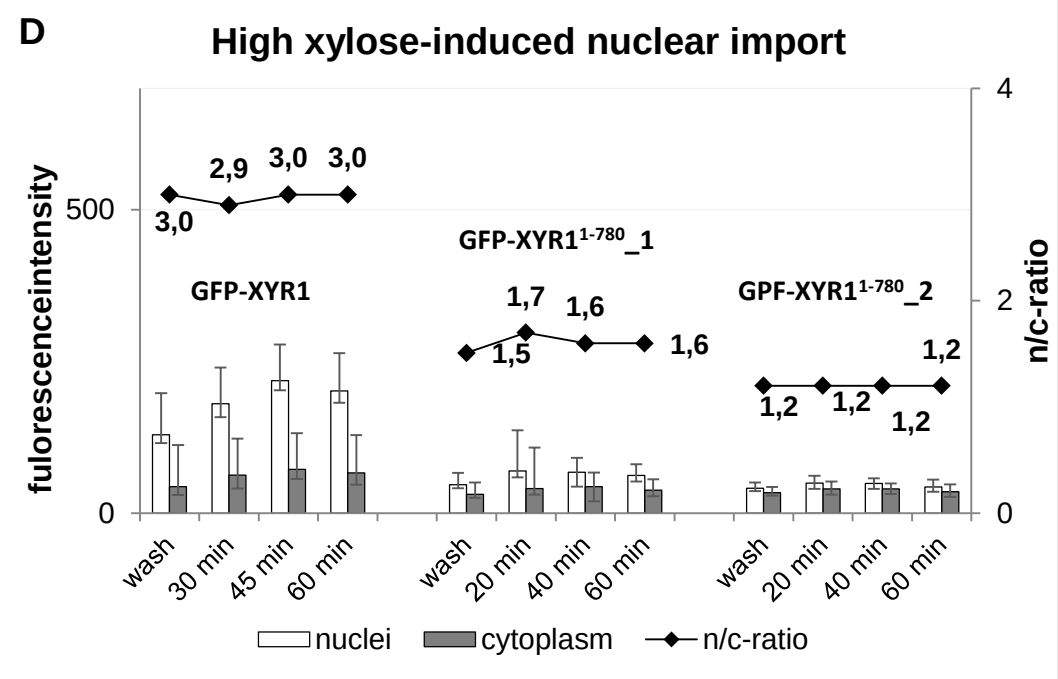


Fig. 15: Nuclear import dynamics of two strains expressing GFP-XYR1$^{1-780}$ compared to a GFP-XYR1 expressing strain in a QM9414 strain background replaced to different carbon sources. **(A)** 1.4mM sophorose. **(B)** 25mM lactose **(C)** 1mM xylose and **(D)** 65mM xylose. Nuclear import of GFP-XYR1 after replacement of a glucose culture to sophorose was followed for 60 min. Fluorescence intensity was measured in the cytoplasm and the nuclei. Representative, inverted fluorescence images of GFP-XYR1 accumulation in nuclei at the indicated time points of the experiment are shown below the graph.

Under all tested cellulase- or xylanase-inducing conditions (25mM lactose, 1.4mM sophorose and 1mM xylose) GFP-XYR1 became recruited into nuclei, as expected, whereas the fluorescence intensity of strains expressing the truncated GFP-XYR1$^{1-780}$ protein did not increase above its baseline level. Sophorose, the strongest known cellulase inducer, results in a faster nuclear import of XYR1 in comparison to 25mM lactose or 1mM xylose . The fluorescence images show that a nuclear recruitment took place and a stronger fluorescent signal of the nuclei is visible. Notably, under high xylose conditions the baseline level of XYR1 protein was elevated in the GFP-XYR1 expressing strain compared to the mutant XYR1 (Fig. 15D). Here, the fluorescent intensity is close to the detection limit of the confocal microscope. Overall, these results demonstrate a lack of increase of the fluorescence intensity in strains expressing GFP-XYR1$^{1-780}$ under different conditions which can be either explained by a rapid degradation of XYR1 or an insufficient transcriptional upregulation of *xyr1* resulting from an auto-regulatory positive feedback mechanism of XYR1 for which the C-terminal putative regulatory domain might be essential.

To ensure that the observed small variations of GFP-XYR1$^{1-780}$ fluorescence truly reflect significant changes in the amount of expressed protein we have eliminated auto-fluorescence

background detection of the imaging set up (Figure A1). This analysis verified that under these conditions GFP-XYR1$^{1-780}$ synthesis and nuclear accumulation found, however, only to an amount comparable to that of GFP-XYR1 under non-inducing conditions, i.e. baseline resting state found under CCR. None of the other carbon sources had a stronger influence to GFP-XYR1$^{1-780}$ when compared to GFP-XYR1. On 65mM xylose, GFP-XYR1$^{1-780}$ showed a repressed nuclear import which was difficult to detect. Under repressing conditions, the measured fluorescent intensity was close to detection limit of the confocal microscope. So the n/c-ratio showed no significant increase. This supports the notion that removal of the putative C-terminal regulatory domain primarily disrupts a signaling function of the auto-regulatory feedback mechanism of XYR1, and not its nuclear import.

3.3 Influence of the XYR1 truncation on key gene expression

In order to correlate nuclear presence of full-length and truncated GFP-XYR1 constructs, respectively, with gene transcription, the influence of XYR1 C-terminal truncation on key gene expression (*xyr1*, *cbh1*, *xyn2* and *cre1*) was determined by RT-qPCR (Fig. 16).

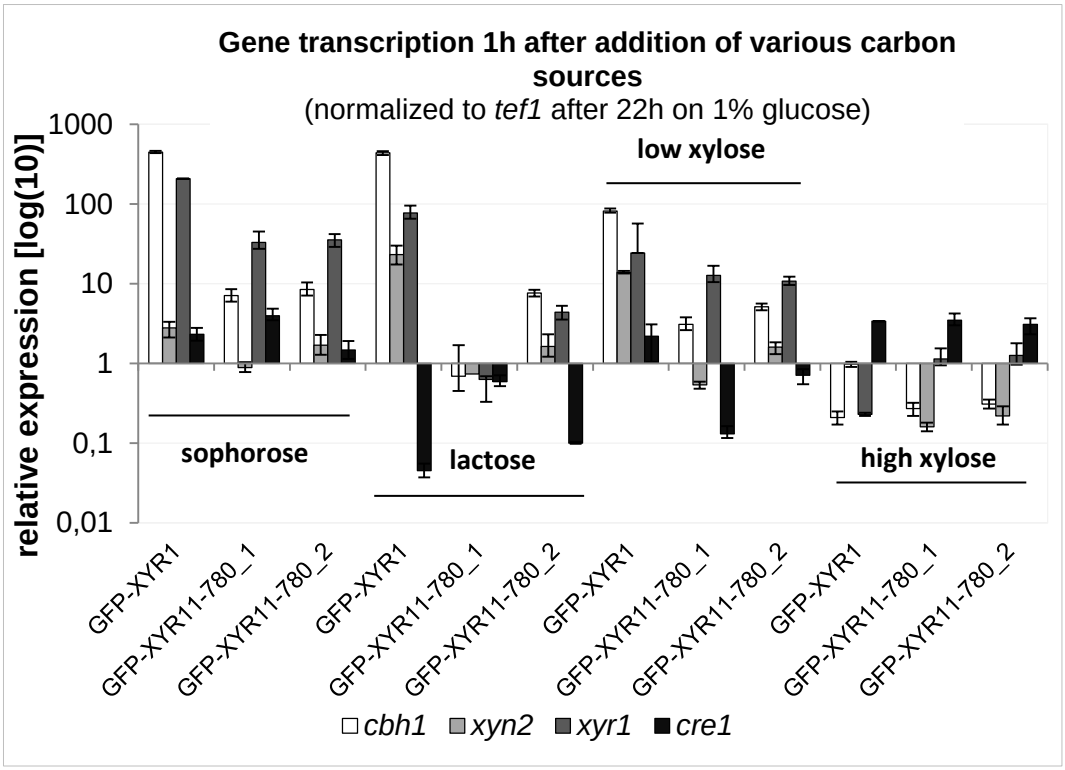


Fig. 16: Gene transcription of GFP-XYR1$^{1-780}$ and GFP-XYR1 expressing QM9414 strains. Transcript levels were determined after 1h of replacement to 1.4mM sophorose, 25mM lactose, 1mM xylose or 65mM xylose after a precultue on glucose as carbon source. All expression data are normalized to *tef1* expression after 22h in 1% glucose.

On cellulase or xylanase inducing carbon sources (1.4 mM sophorose, 1mM xylose and 25 mM lactose), *xyr1* and *cbh1* gene expression were strongly upregulated in the GFP-XYR1 expressing strain TRAL002. In both GFP-XYR1$^{1-780}$ clones, expression of the truncated *xyr1* allele was also clearly upregulated, however, but clearly lower when compared to the control strain TRAL002. More importantly, *cbh1* expression did not significantly respond to the increase in *xyr1*, suggesting that only insufficient amounts of non-functional XYR1 were produced. Alternatively, the low-level *cbh1* expression might result from derepression under these conditions. Compared to 1mM xylose overall expression of the key genes was lower, but the general trends between transformants was similar compared to sophorose, excluding a higher *xyn2* response, as expected with xylose as inducing carbon source. Interestingly, on 25mM lactose only one GFP-XYR1$^{1-780}$ strain showed a positive transcriptional activity of *cbh1*, *xyn2* or *xyr1*, whereas TRAL002 responded as expected with high, correlated *xyr1* and *cbh1* expression. *Cre1* expression was, as expected, consistently low or downregulated under all conditions. With 65mM xylose as the sole carbon source, the transcription profiles were essentially inversed, with all three strains responding with upregulated *cre1* expression to the repressing condition.

The gene expression results combined with the intracellular shuttling dynamic (Fig. 16) demonstrate that C-terminal truncation results in a lower XYR1 gene expression and therefor also a lower nuclear accumulation under inducing conditions. This suggests that C-terminal truncation of XYR1 is involved in a positive feedback auto-regulation of the transcription factor XYR1 and is needed for cellulase and xylanase production

3.4 Colony recruitment profile

Mature colonies of filamentous fungi are functionally stratified into three major zones with distinct morphologies: (i) the colony periphery which comprises fast growing leading hyphae that branch sparsely, and do not differentiate aerial structures; (ii) the colony subperiphery displays prolific hyphal branching, but also does not form aerial hyphae , and (iii) the central area of the colony which is characterized by intense hyphal branching, aerial hyphae formation and prolific sporulation (Figure 17).

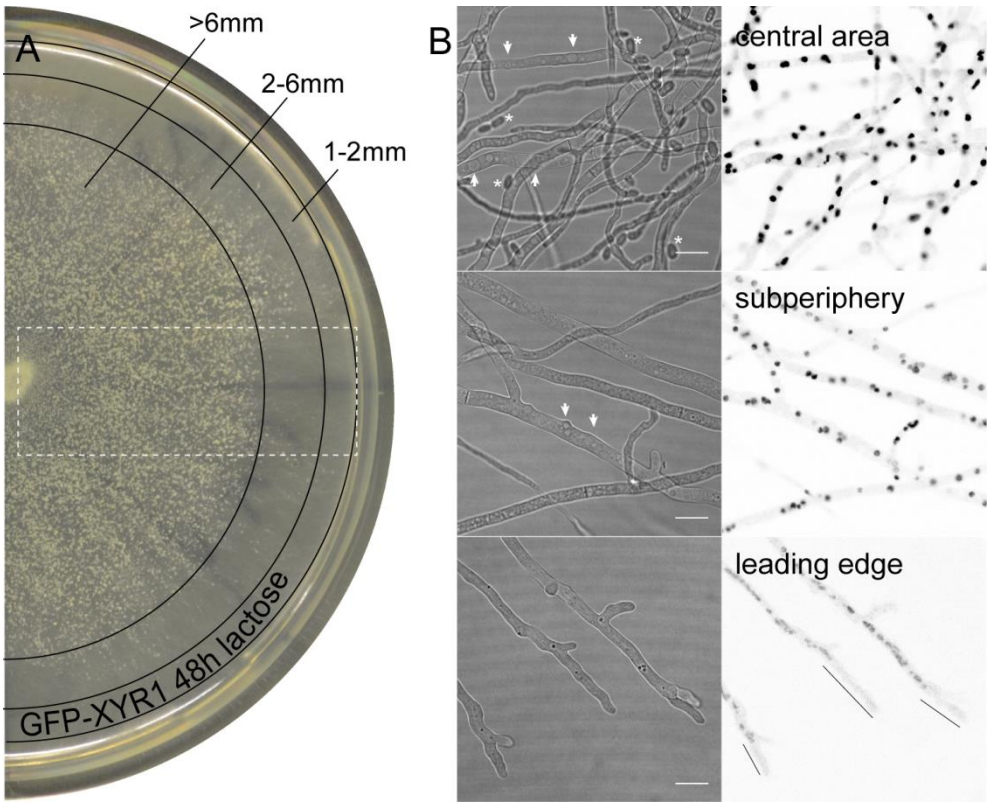


Fig. 17: Transcription factor recruitment in the morphologically and functionally stratified mature fungal colony. **(A)** 48h lactose plate culture of the GFP-XYR1 expressing TRAL002 strain is visible. Circeled areas indicate the three main morphologically distinct zones: periphery, 1-2mm from leading edge; subperiphery, 2-6mm form leading edge; central area, >6mm form leading edge of the colony. The white dotted rectangle indicates the sampling area for live-cell imaging microscopy comprising all three zone.. **(B)** Bright field and inverted fluorescence images showing the typical hyphal phenotypes in the three zones: dense hyphal 'network' with conidiation (conidia indicated with asterisks) and a high degree of vacuolation (vacuoles indicated with arrowheads) in the central area, branched hyphae with beginning vacuolation but no conidiation in the subperiphery, and largely non-vacuolated leading hyphae with clear apical nuclear-exclusion zone (indicated with black lines in the fluorescence image). Modified from (Lichius et al. 2014).

Previous data showed that nuclear recruitment of XYR1 occurs at different degrees in these three functional zones of the colony, and depending on the used carbon source (Lichius et al., 2014. subm.). We therefore, asked the question whether C-terminal truncation might influence the nuclear recruitment dynamics of XYR1 in these different zones

Strains expressing the full-length GFP-XYR1 behaved as expected: strong nuclear recruitment of XYR1 with a clear intensity peek in the colony center on inducing carbon sources and essentially the same profile but with about 10-times weaker peak intensities on non-inducing carbon sources (Fig. 18). The strain expressing the truncated GFP-XYR1$^{1-780}$, on the other hand, showed low-intensity profiles independent of the used carbon source. Notably, the overall amounts of GFP- XYR1$^{1-780}$ on any carbon source were comparable to those of GFP-XYR1 under

non-inducing conditions, suggesting that the induction-independent baseline production GFP-XYR1$^{1-780}$ functioned normally. This is in line with data collected in previous shuttling experiments using conidial germlings in liquid culture (e.g. Figure 16). Taken together these data suggest, that C-terminal truncation of XYR1 prevents any XYR1-dependent cellular response, independent of the used carbon source, the developmental age of the fungus, i.e. germling vs. mature colony, or the functionally stratified colony region.

The colony recruitment data confirms our previous notion suggesting that the putative regulatory C-terminus of XYR1 is not *per se* required for nuclear import of the transcription factor, but rather seems to be involved in its induction-dependent, positive auto-regulatory feedback mechanism. This explains why the fluorescent intensity of the truncated XYR1 protein is generally low because the protein cannot be upregulate by its own production and achieve sufficient nuclear accumulation to trigger cellulase gene expression

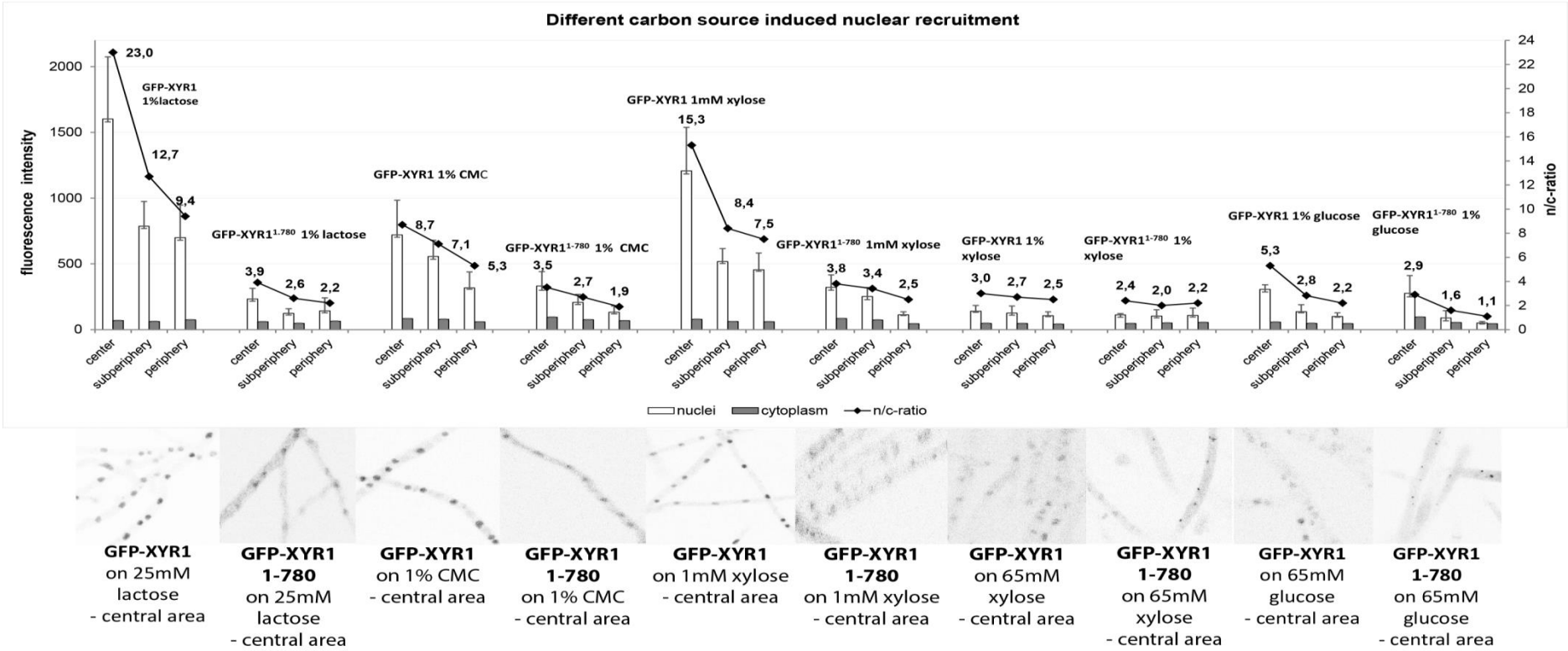


Fig. 18: Colony recruitment profile of full-length and truncated GFP-XYR1. (A) Quantification of nuclear recruitment of GFP-XYR1 and GFP-XYR1$^{1-780}$, respectively, in the three main colony zones. Full-length GFP-XYR1 responds to a cellulase-inducing signal (1% lactose) as expected with intense nuclear recruitment, most prominently in the central area of the colony. Under non-inducing/repressing conditions (1% glucose) only weak baseline nuclear recruitment can be detected. In contrast, the truncated GFP-XYR1$^{1-780}$ does not significantly respond to cellulase induction and hence can only be detected at baseline levels independent of the provided carbon source. **(B)** Representative fluorescence images of the tested conditions, showing that GFP-XYR1$^{1-780}$ does respond to 1% lactose with nuclear import, however, only accumulates in the nucleus at insufficient levels to trigger further cellular processes. On 1% glucose GFP-XYR1$^{1-780}$ localizes to small spots, suggesting degradation of the non-functional XYR1 construct in the proteasome.

4 Discussion

T. reesei is widely used for the industrial production of cellulolytic and hemicellulolytic enzymes and has become a paradigm for research on the regulation of the formation of these enzymes (Xu et al. 2009). With the availability of the *T. reesei* QM6a genome (Martinez et al. 2008) and many details about different aspects of regulation of their expression elaborated in genome wide expression studies (Kubicek 2013), new strategies were proposed to improve our understanding of the outstanding protein expression capacities in this fungus. Beside the improved cellulase producers of the NATICK and RUTGERS series, the non-cellulase producing strains such as QM9136 could also help us to find new targets for strain improvement or to lead to new insights in cellulase regulation.

In this master thesis the genetic basis of the cellulase negative strain QM9136 was investigated and the work shed new light on the role of the putative regulatory C-terminal domain of XYR1 in *T.reesei*. In *A. niger,* this domain was described to be involved in nuclear import and transcriptional regulation during cellulase and xylanase induction (Hasper et al. 2004). In this study the scientific questions whether the C-terminal truncation of XYR1 is responsible for the cellulase negative phenotype of QM9136, whether this is the only mutation found in QM9136 leading to the cellulase negative phenotype and how this truncation influences nuclear import and cellulase and xylanase gene expression, were pursued.

The first direct evidence that the XYR1 truncation is responsible for the cellulase negative phenotype came from the reintroduction of the native *xyr1* in *T. reesei*. Both growth test on xylose as well as transcriptional analysis of *cbh1*, *xyr1* or *xyn2* showed that the parental phenotype was restored and that the mutation of *xyr1* was solely responsible for this cellulase negative phenotype. Similar to the knock out strain Δ*xyr1*, QM9136 is severely impaired to grow on D-xylose. XYR1 regulates the D-xylose reductase gene *xyl1* and other genes of the xylose pathway (B. Seiboth, unpublished data) and this is therefore the reasons for the compact growth phenotype of the Δ*xyr1* on higher concentrations of D-xylose. Similarly the D-xylose reductase transcription seems to be also affected by the truncation of the XYR1 in QM9136. The truncated form of XYR1 itself seems to result in no obvious phenotype, as an introduction of the mutated form in the Δ*xyr1* strain did not lead to a cellulase positive phenotype. It might have still a small activating function as the levels of the cellulase and xylanase genes were a bit higher under inducing conditions but this will need further experimental support to clarify this assumption.

Strains expressing the truncated form of XYR1 fused to GFP in a QM9414 background resulted in a cellulase negative phenotype and a strong growth retardation on xylose supporting the conclusions discussed above and excluding that GFP would have an effect on gene expression. However, when the mutated copy of *xyr1* is replaced by a *gfp-xyr1* fusion, the cellulase negative phenotype is reversed, demonstrating that the GFP-XYR1 fusion is fully functional and can also rescue the QM9136 loss-of-function phenotype underlining that the mutation found in *xyr1* is the only responsible for the cellulase negative phenotype. This is in line with previous complementation studies where GFP did not influence the behavior of XYR1 (Lichius et al, 2014). Thus, the GFP-XYR1 expressing strains can be used as a fully functional reference strain for GFP-XYR1$^{1-780}$ during this work.

The comparison of a GFP-XYR1 to a GFP-XYR1$^{1-780}$ gives now also a first functional explanation for the defect found in QM9136 and highlight the function of the XYR1 C-terminus.

As a prerequisite of this analysis it was also shown that the strain background does not influence shuttling of GFP-XYR1 and that this fusion works equally well in both strain backgrounds of QM9414 and QM9136. GFP-XYR1 can therefore fully compensate XYR1 loss-of-function and thus the phenotype of QM9136.

Under inducing and non-inducing conditions GFP-XYR1$^{1-780}$ nuclear localization was reduced and remained on a low level. Under non-inducing conditions the fluorescence intensity of GFP-XYR1$^{1-780}$ was even close to the detection level of the confocal microscope. Nevertheless, biosynthesis and nuclear import of GFP-XYR1$^{1-780}$ increased slightly under inducing conditions. The best inducer for *xyr1* is sophorose followed by lactose and low xylose concentration. Elimination of the auto-fluorescent background under inducing and non-inducing conditions, GFP- XYR1$^{1-780}$ could be localized as small cellular clusters (spots), and its cytoplasmic fraction was slightly elevated, possibly suggesting less efficient nuclear import. Under non-inducing condition, GFP-XYR1$^{1-780}$ was only weakly detectable in the cytoplasm. This suggests that the level of the truncated XYR1 allele respond to cellulase-induction, but that its accumulation is by far lower than found for the GFP-XYR1 fusion.

The data from recruitment of XYR1 in the stratified colony support these results and suggest that the main function of the XYR1 C-terminus is not nuclear localization that this part is rather important for a XYR1 auto-regulation under inducing conditions.

So the XYR1 truncation with the lack of the C-terminal domain prevents an upregulation of XYR1 and although XYR1 biosynthesis and import occur the truncated XYR1 is not able to stimulate XYR1 expression.

Analysis of the *cbh1* and *xyn2* gene transcription of GFP-XYR1$^{1-780}$ in comparison to GFP-XYR1 showed a basal level of gene expression. In the study of Torigoi et al. 1996, their analysis of the mutant strain QM9136 revealed that this strain is defective in the induction process but not in the basal gene expression of the cellulase genes. The authors concluded this from a strain in which the cellulase gene promoter *cbh1* was fused to the hygromycin selection marker. Our results would be in support of this claim but further experiments would be necessary to compare if our basal expression is similar to the one observed by Torigoi et al. 1996. Other factors influencing the expression of the hygromycin resistance gene could be still found in the different location of the *cbh1-hph* construct in the genome or the length of the promoter region which could influence the level of *hph* expression in this strain. This might be therefore clarified by the replacement of the *hph* gene in the *cbh1* locus.

In summary, several novel insights in cellulase regulation are the result from this work and are highlighted by the finding that the C-terminus is important for the upregulation of XYR1 during inducing conditions, thereby influencing cellulase and xylanase gene expression. Interestingly, the proposed auto-regulatory feedback mechanism seems to be switched off due to the loss of the C-terminal domain of the XYR1$^{1-780}$ and *cbh1* gene expression remains only on a basal level which is slightly increased.

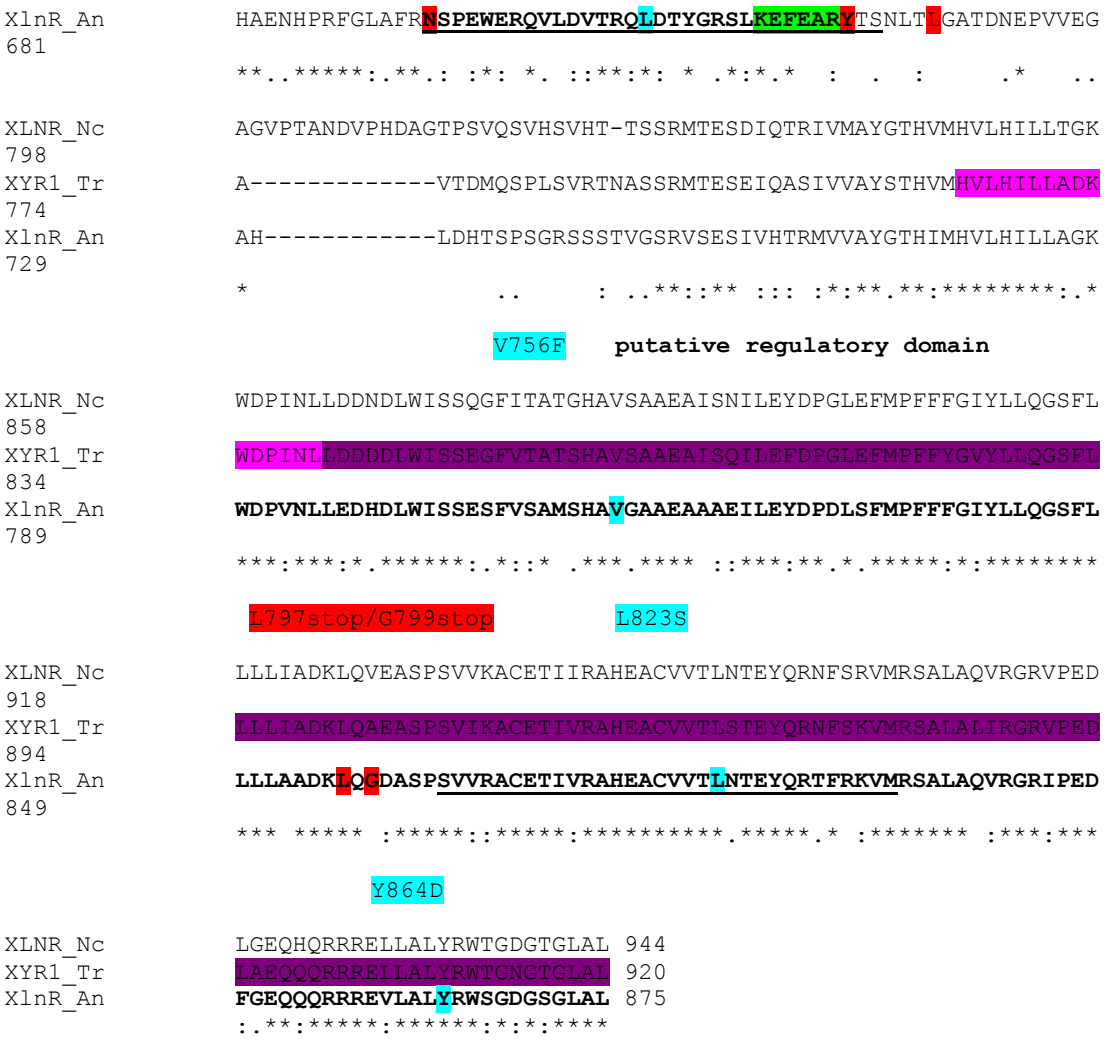
5 Appendix

5.1 Protein sequence and functional domain alignment of filamentous fungal XYR1 - orthologous

Nc = N.crassa, Tr = T. reesei, An = A. niger:

```
XLNR_Nc         MLSNPLHRFAPYHAMPSPTLLSGGHVTASHLHAAGLDTMGPGSHYALQQLQQHVSVHNHH
60
XYR1_Tr         MLSNPLRRYSAYPDISSASFDPNYHGSQSHLHSINVNTFGNSHPYPMQHLAQHAELSSSR
60
XlnR_An         ----------------------MSHTKDQPPFDNEKNQSTGSGFRDALQ---------RDP
30
                                     .     . :.    :: * .   .:*

                                                GAL4-like Zn(II)2Cys6 DNA binding

XLNR_Nc         LARAGP-QPKHRQHPYGPVTR-ATGAAGPIRRRISRACDQCNQLRTKCDGQHPCAHCIEF
118
XYR1_Tr         MIRASPVQPKQRQGSLIAARKNSTGTAGPIRRRISRACDQCNQLRTKCDGLHPCAHCIEF
120
XlnR_An         LVEARS-----------AVRK--TSSSAPVRRRISRACDQCNQLRTKCDGQHPCAHCIEF
77
                : .* .           .. :  *.::.*:******************** *********

                 Y83stop  putative quadruplepartite NLS → conserved tripartite
NLS

XLNR_Nc         GLGCEYIRERKKRGKASRKDLAAQAAAAAAAQ--LNGHKNPSQAGENDQSPPNRT-----
171
XYR1_Tr         GLGCEYVRERKKRGKASRKDIAAQQAAAAAAQ--HSGQVQDGPEDQHRKLSRQQS-----
173
XlnR_An         GLTCEYARERKKRGKASKKDLAAAAAAATQGSNGHSGQANASLMGERTSEDSRPGQDVNG
137
                ** *** **********:**:**  ***: ..   .*: : .  .:. .   .

XLNR_Nc         --ESTTATKRASSLPIEHQTTSNDKTMSDMSE-GSVRSQRTGSMDSIDLGAHQTHIASHP
228
XYR1_Tr         --ESSRGSAELAQPAHDPPHGHIEGSVSSFSDNGLSQHAAMGGMD--GLEDHHGHVGVDP
229
XlnR_An         TYDSAFESHHLSSQPSHMQHASTAG-ISGLHESQTAPSHSQSSLGTTIDAMHLNHFNTMN
196
                  :*:  : . :. . .         :*.: :         ..:.      *  *.

XLNR_Nc         GAMDRDLESPAALDL-SYGNVHQEYHR-QGMGAHLMNGASHHTPYGSNQAAMSNYPDLPY
286
XYR1_Tr         ALGRTQLEASSAMGLGAYGEVHPGYES-PGMNGHVMVPPS----YG-AQTTMAGYSGISY
283
XlnR_An         DSGRPAMSISDLRSLPPSVLPPQGLSSGYNASAFALVNPQEPGSPANQFRLGSSAENPTA
256
                      :. .   .* .            . ... :  ..     .      :.  . .

XLNR_Nc         ALHTQSPTGYSANTSSGFRIGASPLSAYPMAGGSTSPGWMNLASPPPQFAQHIPQPTYSH
346
```

```
XYR1_Tr         AAQAPSPATYSS--DGNFRLTG-HIHDYPLANGSS--------------------PSWGQ
320
XlnR_An         PFLGLSPPGQSP---GWLPLPSPSPANFPSFSLHP----------------------FS
290
                .    **.  *.   . : : .     :*  .  .

XLNR_Nc         AQLRYPVLEPLLPHLGNLMPVSLACDLIDLYFASSSSAQMHPMSPYVLGFVFRKRSFLHP
406
XYR1_Tr         SDLRYPVLEPLLPHLGNILPVSLACDLIDLYFSSSSSAQMHPMSPYVLGFVFRKRSFLHP
380
XlnR_An         STLRYPVLQPVLPHIASIIPQSLACDLLDVYFTSSSSSHLSPLSPYVVGYIFRKQSFLHP
350
                : ******:*:***:..::* ******:*:**:****::: *:****:*::***:*****

XLNR_Nc         TKPRQCQPALLASMLWVAAQTSDAPFLTSVPSARGKICQKLLELTVSLLKPLIH---TPS
463
XYR1_Tr         TNPRRCQPALLASMLWVAAQTSEASFLTSLPSARSKVCQKLLELTVGLLQPLIH---TGT
437
XlnR_An         TKPRICSPGLLASMLWVAAQTSEAAFLTSPPSARGRVCQKLLELTIGLLRPLVHGPATGE
410
                *:** *.*.*************:*.**** ****.::********:.**:**:*   *

XLNR_Nc         EEPSPVSSPIVDGVALGGLGVALPGSISMDALTGETGAFGAAGTLDDVVTYIHLATVVSA
523
XYR1_Tr         NSPSPKTSPVVGAAALGVLGVAMPGSLNMDSLAGETGAFGAIGSLDDVITYVHLATVVSA
497
XlnR_An         ASPNYAANMVINGVALGGFGVS------MDQLGAQS---SATGAVDDVATYVHLATVVSA
461
                 .*.  :. ::...*** :**:      ** * .::   .* *::*** **:********

                 fungal-specific transcription factor domain (highly
conserved)

XLNR_Nc         SEYKGASLRWWNAAWSLARELKLGREIPQ-NSPSMQNSGSELDGEMG-------------
569
XYR1_Tr         SEYKGASLRWWGAAWSLARELKLGRELPPGNPPANQEDGEGLSEDVDEHDL---------
548
XlnR_An         SEYKAASMRWWTAAWSLARELKLGRELPPNVSHARQDGERDGDGEADKRHPPTLITSLGH
521
                ****.**:*** **************:*   . : *:.    . : .

XLNR_Nc         ---NIPGMITEEEREERRRIWWLVYIVDRHLALCYNRPLFLLDIECDGLLQPMDDTDYQN
626
XYR1_Tr         -NRNNTRFVTEEEREERRRAWWLVYIVDRHLALCYNRPLFLLDSECSDLYHPMDDIKWQA
607
XlnR_An         GSGSSGINVTEEEREERRRLWWLLYATDRHLALCYNRPLTLLDKECGGLLQPMNDDLWQV
581
                   .    :********** ***:* .************ *** **..* :**:*  :*


XLNR_Nc         GNFYAYTDPN-------VLASDPNTPAARHRGPSFVCTGHSIFGYFLPLMTILGEIVDLH
679
XYR1_Tr         GKFRSHDAGNSSINIDSSMTDEFGDSPRAARGAHYECRGRSIFGYFLSLMTILGEIVDVH
667
XlnR_An         G--------------------DFAAAAYRQVGPPVECTGHSMYGYFLPLMTILGGIVDLH
621
                *                    :   ..    *.   * *:*::****.****** ***:*

                         putative coiled-coil/nuclear import domain(?)

                         N635stop   Leu650Pro   Y664stop   L668stop
                                          tyrosine (Y)phosphorylation site

XLNR_Nc         HARNHPRFGVGFRSSREWDDQTAEITRHLEIYEESIKRFEHRNLSLSAQAQAADEKAAEA
739
XYR1_Tr         HAKSHPRFGVGFRSARDWDEQVAEITRHLDMYEESLKRFVAKHLPLSSKDKEQHEMHDSG
727
```


XlnR_An 681 HAENHPRFGLAFRNSPEWERQVLDVTRQLDTYGRSLKEFEARYTSNLTLGATDNEPVVEG
XLNR_Nc 798 AGVPTANDVPHDAGTPSVQSVHSVHT-TSSRMTESDIQTRIVMAYGTHVMHVLHILLTGK
XYR1_Tr 774 A----------VTDMQSPLSVRTNASSRMTESEIQASIVVAYSTHVMHVLHILLADK
XlnR_An 729 AH----------LDHTSPSGRSSSTVGSRVSESIVHTRMVVAYGTHIMHVLHILLAGK
V756F
putative regulatory domain
XLNR_Nc 858 WDPINLLDDNDLWISSQGFITATGHAVSAAEAISNILEYDPGLEFMPFFFGIYLLQGSFL
XYR1_Tr 834 WDPINLLDDDDLWISSEGFVTATSHAVSAAEAISQILEFDPGLEFMPFFYGVYLLQGSFL
XlnR_An 789 WDPVNLLEDHDLWISSESFVSAMSHAVGAAEAAAEILEYDPDLSFMPFFFGIYLLQGSFL
L797stop/G799stop
L823S
XLNR_Nc 918 LLLIADKLQVEASPSVVKACETIIRAHEACVVTLNTEYQRNFSRVMRSALAQVRGRVPED
XYR1_Tr 894 LLLIADKLQAEASPSVIKACETIVRAHEACVVTLSTEYQRNFSKVMRSALALIRGRVPED
XlnR_An 849 LLLAADKLQGDASPSVVRACETIVRAHEACVVTLNTEYQRTFRKVMRSALAQVRGRIPED
Y864D
XLNR_Nc LGEQHQRRRELLALYRWTGDGTGLAL 944
XYR1_Tr LAEQQQRRRELLALYRWTGNGTGLAL 920
XlnR_An FGEQQQRRREVLALYRWSGDGSGLAL 875

5.1.1 Elimination of the Auto-fluorescent Background

To verify that the generally weak GFP-XYR1$^{1-780}$ fluorescence signals were no artifacts, microscope settings were optimized to eliminate cellular background auto-fluorescence, so that any fluorescence signal must emanate from a GFP source. For this test we deliberately choose a weak inducer (1mM xylose) to evaluate GFP-XYR1$^{1-780}$ fluorescence at the lower end of the detection range (Figure A1). The truncated GFP-XYR1$^{1-780}$ protein was detectable with an average intensity of about 4.5-times less compared to the full-length GFP-XYR1 construct. Interestingly, GFP-XYR1$^{1-780}$ was localized as small cellular clusters (spots), and its cytoplasmic fraction was slightly elevated, possibly suggesting less efficient nuclear import.

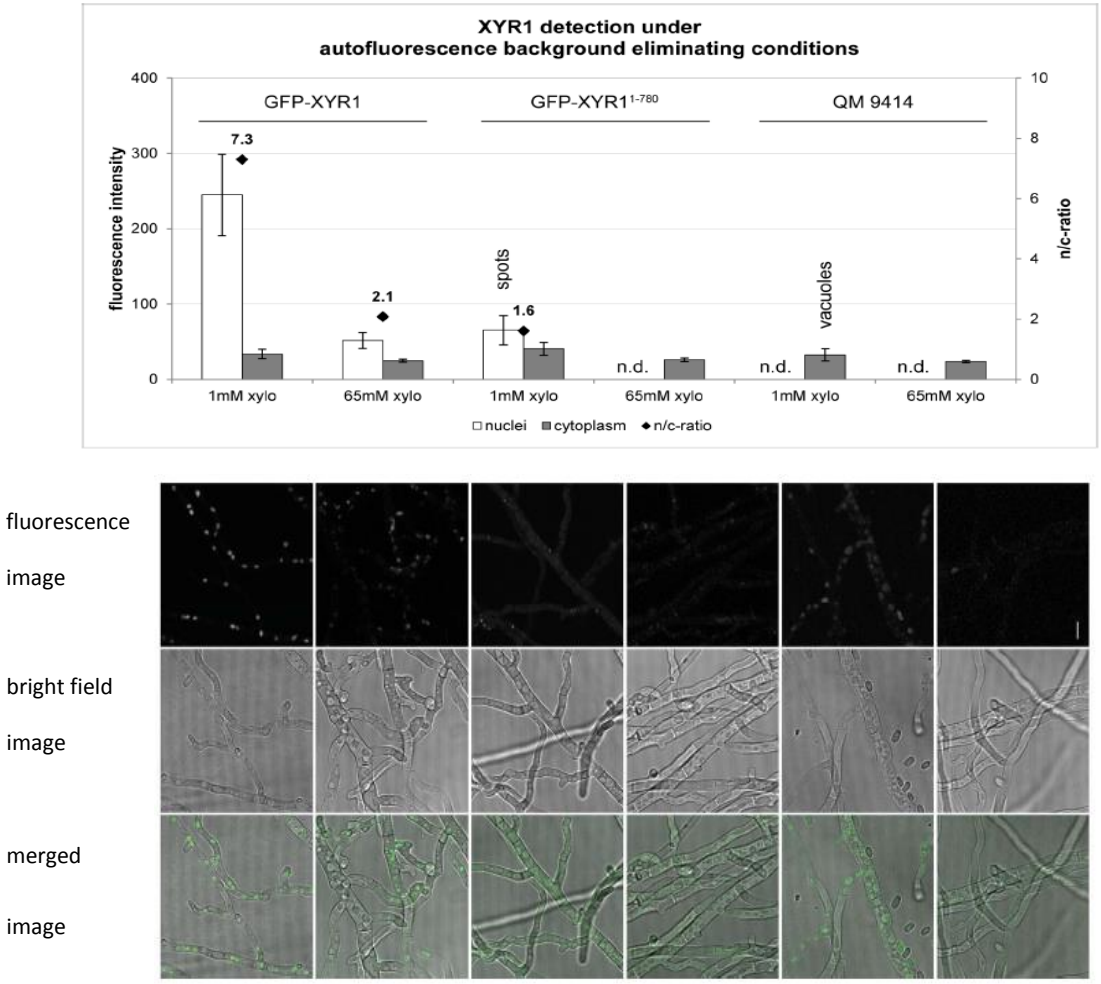


Fig. A1: Elimination of the auto-fluorescent background. Strains are TRAL002 (GFP-XYR1), TRAL007 (GFP-XYR1$^{1-780}$) and QM9414. Each strain was grown on 1mM and 65 mM xylose, respectively, for 48h.

Overall, the level of GFP-XYR1$^{1-780}$ under inducing conditions was within the range of GFP-XYR1 under non-inducing, CCR steady-state, conditions. Under non-inducing condition, GFP-XYR1$^{1-780}$ was only weakly detectable in the cytoplasm. This suggests that the truncated XYR1 allele does respond to cellulase-induction, however, its production cannot be upregulated sufficiently. The

small, bright clusters in which GFP-XYR1$^{1-780}$ localized under these conditions might be degradation vacuoles in which the non-functional construct became quickly turned over, preventing significant nuclear accumulation. Importantly, in the presence of a stronger cellulase-inducing carbon source, such as 1% lactose, the GFP-XYR1$^{1-780}$ construct was detectable in nuclei (see Figure 18).

6 References

Aro, N., M. Ilmen, A. Saloheimo and M. Penttila (2001). "ACE2, a novel transcriptional activator involved in in regulation of cellulase and xylanase expression." FEMS Microbiol. **276**: 24309-24314.

Aro, N., M. Ilmen, A. Saloheimo and M. Penttila (2003). "ACE1 of *Trichoderma reesei* is a repressor of cellulase and xylanase expression." Appl. Environ. Microbiol. **69**: 56-65.

Beals, C. R., C. M. Sheridan, C. W. Turck, P. Gardner and G. R. Crabtree (1997). " Nuclear export of NF-ATc enhanced by glycogen synthase kinase-3." Science **275**: 1930-1934.

Becker, J., F. Melchior, V. Gerke, F. R. Bischoff, H. Ponstingl and A. Wittinghofer (1995). "RNA1 encodes a GTPase-activating protein specific for Gsp1p, the Ran/TC4 homologue of *Saccharomyces cerevisiae.*" J. Biol. Chem. **270**: 1180-11865.

Bedard, J. E., J. D. Purnell and S. M. Ware (2007). "Nuclear import and export signals are essential for proper cellular trafficking and function of ZIC3." Human Molecular Genetics **16**: 187-192.

Blank, T. E., M. P. Woods, C. M. Lebo, P. Xin and J. E. Hopper (1997). "Novel Gal3 proteins showing altered Gal80p binding cause constitutive transcription of GAL4p-activated genes in *Saccharomyces cerevisiae.*" Mol Cell Biol **17**: 2566-2575.

Cherry, J. R. and A. L. Fidantsef (2003). "Directed evolution of industrial enzymes: an update. ." Current Opinion in Biotechnology **14**: 438-443.

Clark, K., M. Ohtsubo, T. Nishimoto, M. Goebl and G. F. Sprague (1991). "The yeast SRM1 protein and human RCC1 protein share analogous functions." Cell. Reg. **2**: 781-792.

Durand, H., M. Clanet and G. Tiraby (1988). "Genetic improvement of Trichoderma reesei for large scale cellulase production." Enzyme Microbiol.Technol. **10**: 341-346.

Eveleigh, D. E. and B. S. Montenecourt (1979). "Increasing yields of extracellular enzymes." Adv. Appl. Microbiol. Mol. Biol. Rev. **25**: 57-74.

Fernandes- Martinez, J., C. V. Brown, E. Diez, J. Tilburn, H. N. Arst, M. A. Penalva and E. A. Espeso (2003). "Overlap of nuclear localization signal and specific DNA-binding residues within the zinc finger domain of PacC." J. Mol. Biol. **334**: 667-684.

Fornerod, M., M. Ohono, M. Yoshida and I. W. Mattaj (1997). "CRM1 is an export receptor for leucin-rich nuclear export signals." Cell. **90**: 1051-1060.

Fukuda, M., S. Asano, T. Nakamura, M. Adachi and M. Yoshida (1997). "CRM1 is responsable for intracellular transport mediated by nuclear export signals." Nature **390**: 308-311.

Görlich, D. (1998). "Transport into and out of the cell nucleus." EMBO J. **17**: 767-778.

Görlich, D., S. Kostka, R. Kraft, S. Dingwall, R. A. Laskey, E. Hartmann and S. Prehn (1995). "Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope." Curr. Biol. **5**: 383-392.

Görlich, D., U. D. Laemmli, Y. Adachi and M. Kohler (1999). "Evidence for distinct substrate specificities of importin alpha family members in nuclear protein import." EMBO J. **18**: 4348-4358.

Gorner, W., E. Durchschlag, M. T. Martinez-Pastor, F. Estruch, G. Ammerer, B. Hamilton, H. Ruis and C. Schuller (1998). "Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity." Genes Dev. **12**: 586-597.

Hasper, A., L. M. Trindade, D. van der Veen, A. L. van Ooyen and L. H. de Graaf (2004). "Functional analysis of the transcriptional activator XlnR from *Aspergillus niger.*" Microbiology **150**(1367-1375).

Herold, S., R. Bischof, B. Metz, B. Seiboth and C. P. Kubicek (2013). "Xylanase gene transcription in *Trichoderma reesei* is triggerd by different inducers representing differnet hemicellulosic pentose polymers." Eukaryontic Cell **12**(3): 390-398.

Hickey, P. C., S. R. Swift, M. G. Roca and N. D. Read (2005). "Live-cell Imaging of filamentous fungi using vital fluorescent dyes and confocal microscopy. In: Methods in Microbiology." Elsevier B.V: 63-87.

Hopper, A. K., H. M. Traglia and R. W. Dunst (1990). "The yeast RNA1 gene product necessary for RNA processing is located in the cytosol and apparently excluded from the nucleus." J. Cell. Biol. **111**: 309-321.

Horn, S. J., M. Sørlie, K. M. Vårum, P. Väljamäe and V. G. H. Eijsink (2012). "Measuring processivity." Methods in enzymology **510**: 69-95.

Horn, S. J., G. Vaaje-Kolstad, B. Westereng and V. G. H. Eijsink (2012b). "Novel enzymes for the degradation of cellulose." Biotechnol. Biofuels **5**: 45.

Ilmén, M., M. L. Onnela, S. Klemsdal, S. Keränen and M. Penttilä (1996). "Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus Trichoderma reesei. ." Molecular & general genetics : MGG **253**: 303-314.

Ilmen, M., A. Saloheimo, M. L. Onnela and M. E. Penttilä (1997). "Regulation of cellulase gene expression in the filamentous fungus Trichoderma reesei." Applied and Environmental Microbiology **63**(4): 1296-1306.

Izaurralde, E. and S. Adam (1998). "Transport of macromolecules between the nucleus and the cytoplasm." RNA **4**: 351-364.

Izaurralde, E., U. Kutay, C. von Kobbe, I. W. Mattaj and D. Görlich (1997). "The asymmetric distribution of the constituents of the Ran system is essential for transport into and out of the nucleus." EMBO J. **16**: 6535-6547.

Khandeparker, R. and M. T. Numan (2008). "Bifunctional xylanases and their potential use in biotechnology. ." Journal of industrial microbiology & biotechnology **35**(7): 635–644.

Kubicek, C., M. Mikus, A. Schuster, M. Schmoll and B. Seiboth (2009). "Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina.*" Biotechnol. Biofuels **2**(19).

Kubicek, C. P. (2013). "Systems biological aproaches towards understanding cellulase production by *Trichoderma reesei.*" Journal of Biotechnology **163**: 133-142.

Kubicek, C. P., V. Seidl and B. Seiboth (2010; Wiley). Plant Cell Wall and Chitin Degradation. (396–415): ISBN: 978-1-55581-473-1.

Kubodera, T., N. Yamashita and A. Nishimura (2002). "Transformation of *Aspergillus sp*. and *Trichoderma reesi* using the pyrithiamine resistence gene (*ptr*A) of *Aspergillus oryzae*." Biosci Biotechnol Biochem **66**: 404-406.

LaCasse, E. C. and Y. A. Lefebvre (1995). "Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins." Nucleic Acids. Res. **23**: 1647-1656.

Laity, J. H., B. M. Lee and P. E. Wright (2001). "Zinc finger proteins: new insights into structural and functional diversity." Curr. Opin. Struct. Biol. **11**: 39-46.

Lichius, A., K. M. Lord, C. E. Jeffree, R. Oborny, R. Booyarungsrit and N. D. Read (2012). "Importance of MAP kinases during protoperithecial morphogenesis in *Neurospora crassa*." PLoS One **7**: 1-21.

Lichius, A., V. Seidl-Seiboth, B. Seiboth and C. P. Kubicek (2014). "Nucleo-cytoplasmic shuttling dynamics of the transcriptional regulators XYR1 and CRE1 under conditions of cellulase and xylanase gene expression in Trichoderma reesei." Acta Microbiologica et Immunologica Hungarica **60**: 179.

Lupas, A., M. van Dyke and J. Stock (1991). "Predicting coiled coils from protein sequences." Science **252**: 1162-1164.

Ma, J. and M. Ptashne (1987). "The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80." Cell. **51**: 113-119.

Mach, R. L., J. Strauss, S. Zeilinger, M. Schindler and C. P. Kubicek (1996). "Carbon catabolite repression of xylanase I (xyn1) gene expression in *Trichoderma reesei*." Mol. Microbiol. **21**: 1273–1281.

MacPherson, S., M. Larochelle and B. Turcotte (2006). "A fungal family of transcriptional regulators: the zinc cluster proteins." Microbiol. Mol. Biol. Rev. **70**(3): 583-604.

Mandels, M., J. Weber and R. Parizek (1971). "Enhanced cellulase production by a mutant of *Trichoderma viride*." American Society for Microbiology **21**: 152-154.

Martinez, D., R. M. Berka, b. Henrissat, M. Saloheimo, M. Arvas, S. E. Baker, J. Chapman, O. Chertkov, P. M. Coutinho, D. Cullen, E. G. Dachin, I. V. Grigoriev, P. Harris, M. Jackson, C. P. Kubicek, C. S. Han, I. Ho, L. F. Larrondo, A. L. de Leon, J. K. Magnuson, S. T. Merino, M. Misra, B. Nelson, N. Putnam, b. Robbertse, A. A. Salamov, M. Schmoll, A. Terry, N. Thayer, A. Westerholm-Parvinen, C. L. Schoch, J. Yao, R. Barabote, M. A. Nelson, C. Detter, D. Bruce, C. R. Kuske, G. Xie, P. Richardson, D. S. Rokhsar, S. M. Lucas, E. M. Rubin, N. Dunn-Coleman, M. Ward and T. S. Brettin (2008). "Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina)." Nature biotechnology **26**: 553-560.

Melchior, F., B. Paschal, J. Evans and L. Gerace (1993). "Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor." J. Cell. Biol. **123**: 1649-1659.

Merino, S. T. and C. J. (2007). "Progress and challenges in enzyme development for biomass utilization." Adv. Biochem. Eng. Biotechnol. **108**: 95-120.

Moore, M. S. and G. Blobel (1993). "The GTP-binding protein Ran/TC4 is required for protein import into the nucleus." Nature **365**: 661-663.

Mullis, K. B. (1990). "Target amplification for DNA analysis by the polymerase chain reaction." Ann Biol Clin (Paris). **48**(8): 579-82.

Nakielny, S. and D. Dreyfuss (1999). "Transport of proteins and RNAs in and out of the nucleus." Cell. **99**: 677-690.

Payne, C. M. e. a. (2011). "Multiple functions of aromatic-carbohydrate interactions in a processive cellulase examined with molecular simulation. ." The Journal of biological chemistry **286**: 41028–35.

Peng, G. and J. E. Hopper (2002). "Gene activation by interaction of an inhibitor with a cytoplasmatic signaling protein." Proc Natl Acad Sci U S A **99**: 8548–8553.

Portnoy, T., A. Margeot, R. Linke, L. Atanasova, E. Fekete, E. Sándor, L. Hartl, L. Karaffa, I. Druzhinina, B. Seiboth, S. Le Crom and C. P. Kubicek (2011b). "The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation." BMC Genomics **12**: 269.

Portnoy, T., A. Margeot, V. Seidl- Seiboth, S. Le Crom, F. Ben Chaabane, R. Linke, B. Seiboth and C. P. Kubicek (2011). "Differential regulation of the cellulase transcription factors XYR1, ACE2 and ACE1 in Trichoderma reesei strain producing high and low levels of cellulase." Eukaryontic Cell **10**: 262-271.

Reese, E. T. (1976). "History of the cellulase program at the U.S. army Natick Development Center." Biotechnology and Bioengineering Symposium: 9-20.

Reese, E. T., H. S. Levisons and M. Dowing (1950). "Quartermster culture collection." Farlowia **4**: 45-86.

Sadowski, I., C. Costa and R. Dhanawansa (1996). "Phosphorylation of Gal4p at a single C-terminal residue is necessary for galactose-inducible transcription." Mol Cell Biol **16**: 4879-4887.

Sambrook, J. and D. W. Russell (2006). "The Inoue Method for Preparation and Transformation of Competent E. Coli: “Ultra-Competent” Cells " Cold Spring Harb Protoc: doi:10.1101/pdb.prot3944

Schjerling, P. and S. Holmberg (1996). "Comparative amino acid sequence analysis of the C6 zinc cluster family of the transcriptional regulators." Nucleic Acids. Rev. **24**: 4599-4607.

Schneider-Poetsch, T., J. Ju, D. E. Eyler, Y. Dang, S. Bhat, W. M. Merrick, R. Green, B. Shen and J. O. Liu (2010). "Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin." Nature **6**: 209-217.

Seiboth, B., C. Ivanova and V. Seidl-seiboth (2011). "Trichoderma reesei : A Fungal Enzyme Producer for Cellulosic Biofuels. ." InTech doi:10.5772/16848.

Seiboth, B., C. Ivanova and V. Seidl-seiboth (2011; Wiley). "Trichoderma reesei : A Fungal Enzyme Producer for Cellulosic Biofuels." In Tech: doi:10.5772/16848.

Seidl, V. and B. Seiboth (2010). "*Trichoderma reesei*: genetics approaches to improving strain efficiency." Biofuels **1**(2): 343-354.

Shapouri, H., M. Salassi and N. Fairbanks (2006). "THE ECONOMIC FEASIBILITY OF ETHANOL PRODUCTION FROM SUGAR IN THE UNITED STATES. ." 69.

Stricker, A. R., K. Grosstessner-Hain, E. Würleitner and R. L. Mach (2006). "Xyr1 (Xylansase regulator 1) regulates both the hydrolytic enzyme system and D- xylose metabolism in *Hypocrea jecorina.*" Eukaryontic Cell **5**: 2128-2137.

Stricker, A. R., R. L. Mach and L. H. de Graaff (2008). "Regulation of transcription of cellulases and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*)." Appl. Microbiol. Biotechnol. **78**: 211-220.

Turcotte, B., B. Akache and S. MacPherson (2004). "The zinc cluster proteins: a yeast family of transcriptional regulators." Recent developments in nucleic acid research **1**: 233-249.

Ware, S. M., K. G. Harutyunyan and J. W. Belmont (2006). "Zic3 is critical for early embryonic patterning during gastrulation." Dev. Dyn. **235**: 776-785.

Wolfe, S. A., L. Nekludova and C. L. Pabo (2000). "DNA recognition by Cys(2)His(2) zinc finger proeteins." Annu. Rev. Biophysic. Biomol. Struct. **29**: 183-212.

Xiao, Z., R. Latek and H. F. Lodish (2003). "An extended bipartite nuclear localization signal in Smad4 is required for its nuclear import and transcriptional activity." Oneogene **22**: 1057-1069.

Xu, Q., A. Singh and M. E. Himmel (2009). "Perspectives and new directions for the production of bioethanol using consolidated bioprocessing of lignocellulose." Current Opinion in Biotechnology **20**: 364–371.

Yan, C., E. L. Lee and L. I. Davis (1998). "Crm1p mediates regulated nuclear export of a yeast AP-1-like transcription factor." EMBO J. **17**: 7416-7429.

Zeilinger, S., R. L. Mach, M. Schindler, P. Herzog and C. P. Kubicek (1997). "Different inducibility of expression of the two xylanase genes *xyn1* and *xyn2* in Trichoderma reesei." J. Biol. Chem. **271**: 25624 –25629.

Zeilinger S., M. S., M. Pail, R. L. Mach and C. P. Kubicek (2003). "Nuceosome transactions on the *Hypocrea jecorina* (*Trichoderma reesei*) cellulase promoter *cbh2* associated with cellulase induction." Mol. Genet. Genomics **270**: 46-55.